Mathematical Modelling of Integrated Signalling Networks in Stomatal Guard Cells

A thesis presented for the degree of

Doctor of Philosophy of Imperial College London

by

Mariano Beguerisse Díaz

Department of Life Sciences

Imperial College London

South Kensington Campus

London, SW7 2AZ

March, 2012

I certify that this thesis, and the research to which it refers, are the product of my own work, and that any ideas or quotations from the work of other people, published or otherwise, are fully acknowledged in accordance with the standard referencing practices of the discipline.

Signed:

Copyright

Nothing in life is to be feared, it is only to be understood.
Now is the time to understand more, so that we may fear less.
Maria Skłodowska-Curie

Wir müssen wissen.
Wir werden wissen.
David Hilbert

Abstract

Stomata are tiny pores in plant leaves that regulate gas and water exchange between plants and their environment. Abscisic acid and ethylene are two well-known elicitors of stomatal closure. Yet when stomata are presented with a combination of both signals, they fail to close; these observations are hard to reconcile biologically and their explanation is not easily obtained by experimental means alone. To shed light on this unexplained behaviour, a combination of mathematical, computational, and experimental techniques are used. A differential equation model of stomatal closure is constructed from known biochemical interactions; this modelling process has motivated the collection of experimental measurements of components in the pathway, at time points beyond what is usually found in the literature. The experimental observations include stomatal aperture and hydrogen peroxide production in *Arabidopsis thaliana* guard cells treated with abscisic acid, ethylene, and a combination of both. These measurements show that sustained high levels of hydrogen peroxide are required to achieve stomatal closure and that guard cells exhibit increased antioxidant activity when treated with a combined dose of abscisic acid and ethylene. Additionally, the experimental observations and modelling suggest a distinct role for two antioxidant mechanisms during stomatal closure: a slower, delayed response that is activated by a single stimulus (abscisic acid *or* ethylene) and another more rapid mechanism that is only activated when both stimuli are present. The model indicates that the presence of this 'and' mechanism in the antioxidant response is crucial to explain the lack of closure under a combined stimulus.

Estimating parameters from data is a key stage of the modelling process, particularly in models of biological systems many parameters need to be estimated from sparse and noisy data sets, such is the case of the stomatal closure model presented here. Over the years, a variety of heuristics have been proposed to solve this complex optimisation problem, with good results in some cases yet with limitations in the biological setting. In this thesis, an algorithm for model parameter fitting is developed combining ideas from evolutionary algorithms, sequential Monte Carlo methods and direct search optimisation. The method is shown to perform well even when the order of magnitude and/or the range of the parameters is unknown. The method refines iteratively a sequence of parameter distributions through local optimisation combined with partial resampling from a historical prior defined over the support of all previous iterations. The method is tested on biological models using

both simulated and real experimental data, and it estimates the parameters efficiently even in the absence of *a priori* knowledge about the parameters. Then, this method is used to find the parameter values of the much larger stomatal closure model from experimental observations.

A classic model of linear activation cascades is studied in this thesis. In a special but important case the output of an entire cascade can be represented analytically as a function of the input and a lower incomplete gamma function. If the inactivation rate of any component is altered, the change induced at the output is independent of the position in the cascade of the modified component. These analytical results show how one can reduce the number of equations and parameters in ODE models of cell signalling cascades, and how delay differential equation models can sometimes be approximated through the use of simple expressions involving the incomplete gamma function. The expressions with the lower incomplete gamma functions are used in the construction of the model of stomatal closure to represent the activation dynamics of the antioxidant mechanisms in guard cells. Fitting the stomatal closure model parameters to the data indicates that the antioxidant responses should have two different timescales, which can explain the lack of closure under a combined ABA and ethylene stimulus.

Acknowledgements

I am most grateful to my supervisors Radhika Desikan, Mauricio Barahona, and the late Jaroslav Stark for agreeing to take me as a student, giving me guidance, and having patience while I learnt the basics of biology. Working with you has been a tremendous experience.

On those difficult months when Jaroslav was seriously ill and Radhika was taking care of her new baby, Piers Ingram, Robert Endres, Simon Moon, and John Mansfield went out of their way and gave up their time to make sure I received enough support and help although it was not their duty, thank you so much. I thank Sophia Yaliraki for being so supporting and patient in these last months when this thesis was being written.

I thank the Biotechnology and Biological Sciences Research Council (BBSRC) and Microsoft Research UK for providing the funding for this PhD through the Dorothy Hodgkin Postgraduate Award.

I am grateful to Yvonne Stewart for showing me around the lab and teaching me how to do experiments. I also want to thank Marco Lizzul and Mercedes Hernández Gómez for doing experiments for me.

A number of people have contributed in different ways to my research and to make my time as a student at Imperial College so special: Phil Assheton, Gerardo Aquino, Chris Barnes, Katja Bier, Konstantinos Chatzimichalis, Diana Clausznitzer, Justine Dattani, Emilie Devienne, Elisa Domínguez, Antoine Delmotte, Muna El-Khairi, Elmer Garduño, Sheldon Hall, Ruian Ke, Michał Komorowski, J. Krishnan, Renaud Lambiotte, Nyuk Ling Ma, Kevin Liu, Andrew Mcdonagh, Amy McLennan, Mariale Moreno, Neave O'Clery, Jasmine Pham, Elena Phoka, Mason Porter, Michael Schaub, Katerina Seich Al Basatena, Zahra Sharif, Coralie Simmons, Michael Stumpf, Reiko Tanaka, Zenna Tavares, Pedro Tecedeiro, Sylvain Tollis, Luke Tweedy, Borislav Vangelov, Maria Vigliotti, and Baojun

Wang. I also thank all my friends that from Mexico and elsewhere have given me support and encouragement. Anyone whose name has been wrongly left out of this section is entitled to compensation in the nearest pub or coffee-house.

I could not have completed this project without the support of my parents Mariano and María Elisa, my sisters María José, Maite and Lucía, brothers-in-law David and Santiago. Thank you so much for all the love and support you have given me. Finally, I thank Heather Harrington for being companion and friend, whose support during these years would be difficult to quantify.

London, February 2012.

Mariano Beguerisse Díaz

Table of contents

List of Figures

List of Tables

List of Publications and Submissions

The results presented in this thesis con be found in the following publications and submissions:

- Beguerisse Díaz, M., Wang, B., Desikan, R., and Barahona, M. (2012). Squeeze-and-breathe evolutionary Monte Carlo optimization with local search acceleration and its application to parameter fitting. *Journal of The Royal Society Interface*. doi:10.1098/rsif.2011.0767.
- Beguerisse Díaz, M., Ingram, P. J., Desikan, R., and Barahona, M. (2011). Linear models of activation cascades: analytical solutions and applications. *arXiv:1112.0270*.
- Beguerisse Díaz, M., Hernández Gómez, M. C., Lizzul, A. M., Barahona, M., and Desikan, R. (2012). Compound stress response in stomatal closure: a mathematical model of ABA and ethylene interaction in guard cells. In preparation.
- Harrington, H. A., Komorowski, M., Beguerisse Díaz, M., Ratto, G. M., and Stumpf, M. P. H. (2011). Dependence of MAPK mediated signaling on Erk isoforms and differences in nuclear shuttling. *arXiv:1110.3742*. *Physical Biology*, in press.

Additionally, during the first year of this PhD a significant portion of work and the writing of the following paper was completed:

- Beguerisse Díaz, M., Porter, M. A., and Onnela, J.-P. (2010). Competition for popularity in bipartite networks. *Chaos*, **20**(4), 043101.

Chapter 1

Introduction

1.1 Overview

Plants are important organisms for a variety of reasons, one of them is that ultimately we all depend on them for survival, which makes understanding them a vital task. One remarkable feature of plants is their ability to adapt to the conditions of their surroundings, which allows them to cope and thrive in ever-changing and uncertain environments.

This thesis is concerned with one of the most important mechanisms plants exploit to cope with environmental changes: stomatal closure. Plants posses pores on their leaves known as *stomata*; through these pores plants are able to exchange oxygen and water-vapour for CO_2 in the air. When the exchange of water for CO_2 is not advantageous to the plant, eg when water is scarce, there is not enough light to perform photosynthesis, or under pathogen attack, plants can regulate the aperture of the pores to limit the loss of water until conditions improve. Plants achieve stomatal regulation by sending signals that travel through a complex network of cellular interactions; often, it is a hormonal signal that initiates the process of stomatal closure. This work focuses on two hormones: abscisic acid (ABA) and ethylene, and their interactions in stomata. Both hormones have been widely studied in the literature; ABA is an important hormone involved in seed dormancy and dessication tolerance, and ethylene is a gas hormone involved in fruit ripening, leaf senescence and plant defence against pathogens. Specifically, we are interested in knowing why these two hormones, which on their own cause the stomata to close, fail to do so when

they are present at the same time.

Understanding stomatal closure and why this seemingly contradicting behaviour occurs, invites an interdisciplinary approach. Given the large number of components involved during stomatal closure and the large number of interactions among them, it is difficult to understand the processes by experimental means alone. Using mathematical descriptions of what is understood to occur during stomatal closure, we can keep track of what happens to the components of the system, make hypothesis and predictions that can be tested in the laboratory. New experimental results inform the improvement of the models which are able to make new predictions. The cycle of mathematical-biological cooperation can go on indefinitely and new, fascinating avenues for research (both mathematical and biological) can arise from this process.

1.2 Aims and thesis structure

For this work a combination of experimental, mathematical, and computational approaches are used to understand stomatal closure. The aims of the thesis are:

- The development of a series of mathematical models that help identify key elements from the ABA-ethylene signalling pathway in guard cells that are responsible for the blocking of signals and subsequent lack of stomatal closure.
- Development of mathematical techniques and frameworks to aid the development and analysis of models of stomatal closure.
- Collection of experimental data to validate the models.
- Providing new directions for future research, based on the results obtained by the modelling process and experimental observations.

To meet these aims, a thorough review of the current knowledge about stomatal closure has been made, alongside a review of the mathematical techniques needed to construct models of stomatal closure. The modelling process in this work has highlighted the need for a consistent data-set performed under uniform laboratory conditions, on a single plant species, and for longer times. Data available from published reports are obtained

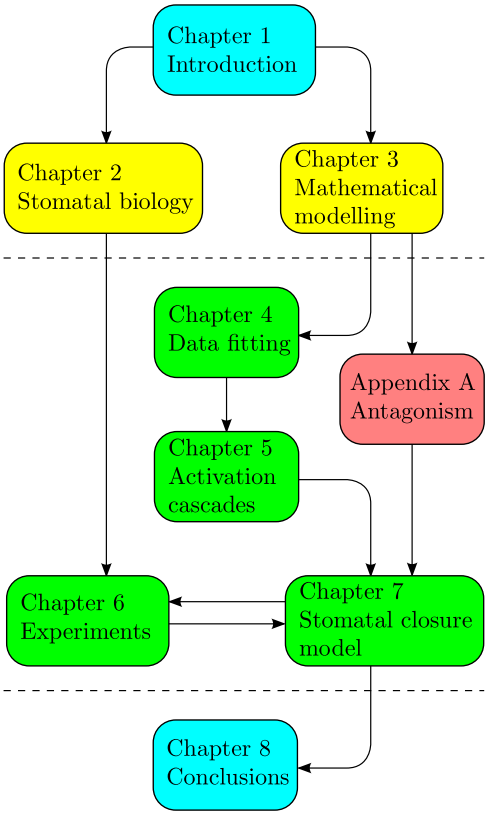

Figure 1.1: Structure of the thesis. In the blue balloons are the introduction and the conclusions of the thesis. Background material is covered in the yellow balloons, and results chapters are in green balloons. Appendix A (in the red balloon) contains additional results.

under experimental conditions which vary considerably in their methodology, measurement times, plant species and cell type. Therefore for this project it was necessary to gather new data from experiments with uniform conditions and time points.

The background information and the results of the thesis are structured by chapters, each dealing with a single topic. The dependencies among the chapters are presented in Fig. 1.1:

- Chapter 2 (background): introduction to stomata, and a thorough review of what is understood to happen after perception of the hormonal signals. The experimental observations that are the motivation for this PhD project are discussed in Sec. 2.5.

- Chapter 3 (background): brief introduction to the mathematical biology techniques

used in this thesis.

- Chapter 4 (results): development of an optimisation method using evolutionary algorithms and Monte Carlo with applications to parameter fitting.
- Chapter 5 (results): analysis of activation cascade models which includes applications to model reduction, equation reordering, and substitution of delay differential equations.
- Chapter 6 (results): experimental observations of guard cell signalling components after treatments with ABA and ethylene.
- Chapter 7 (results): ODE-modelling of stomatal closure. The development of the models is informed by biology in the introductory chapters and the experimental observations, and uses results from Chapters 4 and 5.
- Chapter 8: concluding remarks of the thesis and outlines of what work should be done to take further the research topics addressed in this document.
- Appendix A (additional results): contains explorations of antagonistic toy models and network topologies. This approach has been used to analyse competition among kinases for protein activation in Sec. A.3.

Note that in Fig. 1.1 Chapters 6 and 7 depend on each other. This dependency highlights the iterative nature of the modelling process, where existing data informs models and models prompt new experiments.

An alternative to reading this thesis following the order of the chapters is to read the biological introduction in chapter Chapter 2, followed by the experimental results in Chapter 6, and stomatal modelling in Chapter 7; this ensures that the theme of stomatal closure is not broken by the more mathematical/computational themes in Chapters 3, 4, 5, and Appendix A.

Chapter 2

Biology of stomata

Stomata, guard cells, and their importance for the correct functioning and survival of plants are introduced in this chapter. Then, a description is given of the cellular processes involved in ABA and ethylene-induced stomatal closure, the main focus of this work. Some questions about the signalling processes that still need to be answered are mentioned here, and will be addressed in later chapters.

2.1 Stomata, their function and importance

Unlike animals, plants cannot move to another place whenever a resource (such as water or light) becomes scarce, conditions become unfavourable, or when attacked by predators or disease. Plants possess several physiological mechanisms to help them cope with changes in their surroundings, spread their seeds, fight disease and competitors, and defend themselves from predators.

One of the most remarkable mechanisms of adaptation in plants is the regulation of water vapour and gas exchange. Such regulation responds to diverse stimuli such as changes in light conditions, water-availability, temperature or presence of pathogens. Key players in response to environmental cues are the tiny pores known as *stomata* (singular: *stoma*), located mostly on the lower surface of leaves' epidermis (Freeman, 2008; Taiz and Zeiger, 2002). In the model organism *Arabidopsis thaliana*, each stoma is formed by two kidney-shaped *guard cells*, which are attached to each other by their extremes

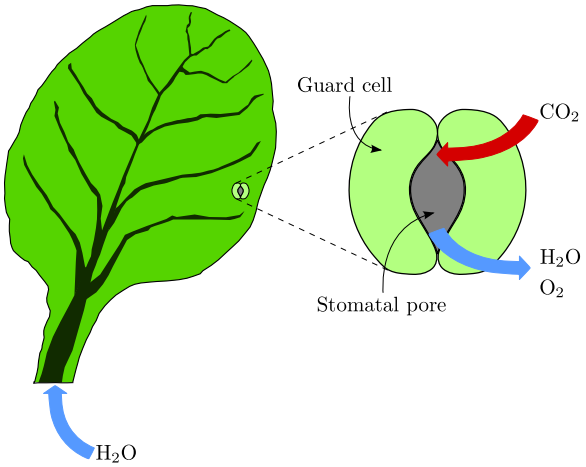

Figure 2.1: Location of stomata on the lower epidermis of a leaf. Water absorbed by the roots is transported into the leaves; if the pore is open, the plant is free to exchange gas and water vapour for CO_2 to perform photosynthesis.

(Fig. 2.1). Although stomata cover only about 5% of the leaf-epidermis area (Hetherington and Woodward, 2003), their importance is difficult to overestimate. Through stomata plants transpire and exchange gases with the atmosphere. The leaf epidermis is almost impermeable to H_2O and other gases, so the main point of gas and water vapour exchange between a plant and the environment is the stomatal pore (Lawson, 2009). When the stomata are open, the plant exchanges H_2O and O_2 for CO_2 from the air, enabling photosynthesis to perform optimally (Acharya and Assmann, 2009). Many environmental factors such as blue-light availability, drought, humidity, CO_2 concentration, or biotic stresses (eg fungi or bacteria) can cause the plant to close its pores, thus slowing its metabolism and minimising water loss (Acharya and Assmann, 2009; Pei *et al.*, 1997; Taiz and Zeiger, 2002).

The pressure of the contents of a plant-cell against its wall, determined by the amount of water contained in the vacuoles, is known as *turgor pressure* (Thoday, 1952). A decrease of the cell's osmotic potential due to ion intake drives water inside the cell and into the vacuoles. When the vacuoles fill with water, the guard cells swell (ie they become *turgid*). Swelling causes the cells to expand in opposite directions but as they are coupled by their extremes and the cell walls are thicker along their height than at the top, this results in the opening of the pore (Fig. 2.2 A). Conversely, when the cell loses water from the vacuoles

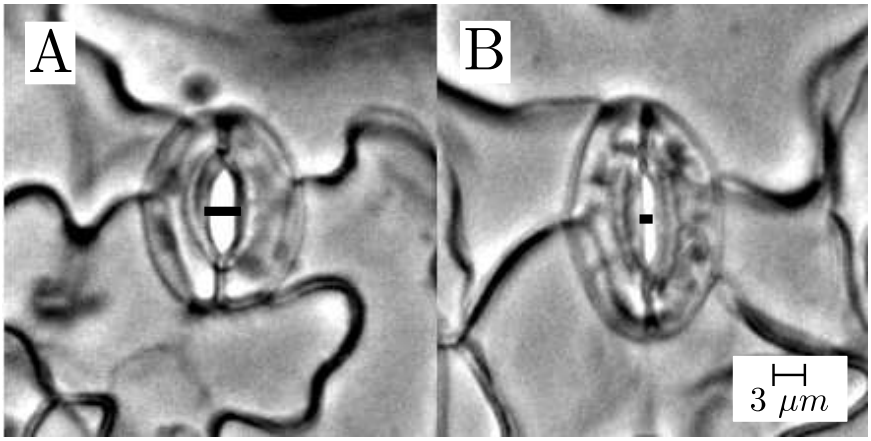

Figure 2.2: Stomata of *Arabidopsis thaliana* leaves under the microscope. **A**: An open stoma, the aperture marked by the bold line is 2.5 μm wide. **B**: A closed stoma, the visibly smaller aperture, compared to the pore on the left, is marked by the bold line across the opening of the pore.

(loss of turgor pressure due to a lower cytosolic ion concentration) their volume decreases, becoming flaccid and ultimately resulting in the closure of the pore (Fig.2.2 B) (Schroeder *et al.*, 2001; Taiz and Zeiger, 2002). Thus, ionic content of the cell is key for controlling cell volume and stomatal aperture. Note also that the shape of guard cells mean that *complete* closure (ie no opening whatsoever) does not occur; a significant reduction in aperture is enough to slow down gas and water exchange to acceptable levels (the exchange never fully stops, as closure is not total).

The regulation of stomatal aperture is at the core of a plant's capacity to cope with changes in its surroundings. Understanding how this process works is important for several reasons:

- Climate change affects the conditions in which many food crops are grown (Gedney *et al.*, 2006). The role of stomata in the planet's water and carbon cycle is a rather important one. Evaporation of water through stomata amounts to 70×10^{15} kg annually (nearly 65% of the precipitation on land). In the tropics alone, stomata release 32×10^{15} kg of water, twice the amount in the atmosphere. In addition, terrestrial photosynthesis fixes 440×10^{15} gCO_2 from the atmosphere annually (Hetherington and Woodward, 2003). Understanding how plants will react to different (perhaps more extreme) environmental conditions is paramount to guaranteeing food security

for all.

- Knowledge of the mechanisms that regulate stomatal movements can inform us of what to expect from plants as the conditions of growth change, and to help engineer and design crops that are more resistant to adverse conditions and are more economical in their need for resources (Li *et al.*, 2006).

- Equally important is that guard cells provide a good framework to study cell-signalling. Hypotheses for possible mechanisms involved in *signal transduction* (ie the relay of an external signal to the inside of the cell to produce a response) can be tested with relative ease by observing whether the stomata are open or closed. The mechanisms of signal transduction in guard cells are likely to be common to many other cell types where experiments are more difficult to perform (Schroeder *et al.*, 2001), which is why it is useful to understand them first in guard cells.

- Finally, guard cell signalling networks are (as this work will attempt to demonstrate) fertile ground for interdisciplinary research where a wide array of approaches from different branches of science (eg genetics, mathematics, computer science, statistics) together with experimental biology have been successfully applied to solve challenging and important problems (Li *et al.*, 2006; Soni *et al.*, 2008).

2.2 Important signalling molecules

This section provides an introduction to some of the most important molecular players involved in the signalling pathways of stomatal closure.

2.2.1 Abscisic acid

Abscisic acid (ABA) is a well-known plant hormone involved in many physiological processes. ABA belongs to the isoprenoid class of metabolites which includes carotenoids, and is produced by most cells containing chloroplasts and amyloplasts (Nambara and Marion-Poll, 2005). It was initially thought that ABA was the hormone responsible for leaf and organ abscission (hence the name abscisic acid), though it was later shown that

abscission is caused by ethylene (Taiz and Zeiger, 2002). ABA was identified early on as the compound responsible for seed dormancy and embryo dessication tolerance (it is sometimes called dormin); ABA-deficient mutants exhibit a wilty phenotype and vivipary (Raz *et al.*, 2001). In the absence of stress, ABA is produced in vascular tissues, and taken up and metabolised by mesophyll cells (Koiwai *et al.*, 2004). Under stress conditions such as drought or salinity, ABA levels in plants increase by biosynthesis and redistribution. Dry roots produce ABA to send a chemical signal to the leaves before drought stress and water deficit are able to cause damage to the plant (Sauter *et al.*, 2001). The ABA signals cause the xylem sap to become more alkaline obstructing mesophyll cell ABA-absorption, and facilitating ABA release which is then able to reach guard cells (Taiz and Zeiger, 2002). In addition to being transported from the roots to the shoots and leaves, ABA is also able to induce its own synthesis (Tan *et al.*, 2003). Inactivation and catabolism of ABA can be done by hydroxylation or conjugation. For example ABA is hydroxylated into phaseic acid (PA) by a cytochrome P450 monooxygenase, and PA levels have been reported to rise after a drought-stressed plant has been re-hydrated, coinciding with a decrease in ABA levels (Nambara and Marion-Poll, 2005). PA is also reported to be physiologically active (Mohr and Schopfer, 1995).

The effect of ABA on stomatal aperture under conditions of drought or salinity is well documented and has been studied extensively in several plant species (Bari and Jones, 2009; Hetherington, 2001; Hubbard *et al.*, 2010; Jones and Mansfield, 1970; Schroeder *et al.*, 2001; Sirichandra *et al.*, 2009; Wasilewska *et al.*, 2008). In Sec 2.3 the signalling events triggered by ABA that lead to loss of turgor and stomatal closure in *Arabidopsis* guard cells are described.

2.2.2 Ethylene

Ethylene is a gaseous hormone involved in several plant processes. In plants ethylene is produced when methionine (Met) is transformed into s-adenosyl-methionine (AdoMet) by AdoMet synthase (as part of the Yang cycle); AdoMet is transformed into the ethylene precursor 1-aminocyclopropane-1-carboxylic acid (ACC) by ACC-synthase; ACC is transformed into ethylene by ACC-oxidase (Taiz and Zeiger, 2002).

There are five known ethylene receptors: ethylene resistant 1 and 2 (ETR1, ETR2), ethylene insensitive 4 (EIN4), ethylene response sensor 1 and 2 (ERS1, ERS2). In the absence of ethylene the receptors activate the protein controlled triple response 1 (CTR1), a negative regulator of the pathway. Active CTR1 inactivates ethylene insensitive 2 (EIN2), a bottleneck in the ethylene pathway below which lie all the other known components and targets. Ethylene inactivates the receptors and CTR1 is unable to inactivate EIN2 and the signal is able to propagate downstream (Chang and Bleecker, 2004; Guo and Ecker, 2004; Stepanova and Alonso, 2005)

The effects of ethylene in plants have been extensively studied in seedlings, and include the control of root elongation, swelling of the hypocotyl and curvature of the apical hook (collectively known as the triple response), fruit ripening, leaf senescence and abscission, defence from pathogens, and seed dormancy (Bari and Jones, 2009; Chang and Bleecker, 2004; Chen *et al.*, 2005; McCue *et al.*, 2009; Taiz and Zeiger, 2002). In some species (such as *Arabidopsis*) ethylene has been found to elicit stomatal closure (Desikan *et al.*, 2006; Pallas and Kays, 1982), whereas in others (eg *Vicia faba*) it is involved in auxin-mediated stomatal opening (Levitt *et al.*, 1987).

2.2.3 Reactive oxygen species

Reactive oxygen species (ROS) are reactive molecules such as superoxide ($O_2^{\bullet -}$), hydrogen peroxide (H_2O_2), hydroxyl radical ($OH^{\bullet}$), and singlet oxygen ($^1O_2$) produced by the incomplete reduction of molecular oxygen (Hancock, 1999). These reactive molecules have in common the ability to cause oxidative damage to DNA, proteins, and lipids, which is why their concentrations in the cell are tightly controlled (Apel and Hirt, 2004). Another common feature among these species (with the possible exception of the hydroxyl radical) is that in sub-toxic concentrations they can function as signalling molecules and second messengers (Cho *et al.*, 2009b; Hancock, 1999; Wang and Song, 2008).

Chemically, ROS are the intermediary species in the process of reducing molecular oxygen to water via a series of single electron transfer steps (Marks *et al.*, 2009):

$$O_2 + e^- \rightarrow O_2^{\bullet -} + e^- \rightarrow H_2O_2 + e^- \rightarrow OH^{\bullet} + e^- \rightarrow H_2O.$$

There are several mechanisms that mediate these reactions, of particular importance for this work are NADPH-oxidases which catalyse the production of superoxide, and superoxide dismutase (SOD) which catalyses the dismutation of superoxide into hydrogen peroxide. ROS are a normal by-product of respiration in many organisms; in plants ROS are the product of photosynthesis and other metabolic processes occurring in chloroplasts, mitochondria and peroxisomes (Neill *et al.*, 2002). Additionally, ROS are produced in response to a variety of stimuli in plants, including abiotic (eg light/dark, ozone (O_3), ultraviolet radiation, and CO_2) and biotic (eg microbial invasion, fungi, and bacteria) (Pham and Desikan, 2009).

As their name suggests ROS are highly reactive, particularly superoxide and the hydroxyl radical. When ROS concentrations are not controlled, cells can suffer from oxidative stress which can lead to impaired physiological function, random damage of the cell machinery, and cell death (Finkel and Holbrook, 2000). Oxidative stress can be a defence mechanism against pathogen invasion, during which up to 15μM H_2O_2 are produced to kill invading microbes, to trigger programmed cell death and limit tissue damage. This production of hydrogen peroxide may be augmented by suppression of antioxidants by salicylic acid and nitric oxide (NO) (Klessig *et al.*, 2000). ROS are also involved in other processes in plants such as plant development, stomatal movements (see below), root growth, and flowering (McInnis *et al.*, 2006; Pham and Desikan, 2009; Zafra *et al.*, 2010).

Plants have a variety of mechanisms to keep ROS levels under control and to reduce them into less-harmful compounds. Non-enzymatic antioxidants include glutathione (GSH), ascorbate (ASC), tocopherol, flavonoids, carotenoids and NO; enzymatic antoxidants include ascorbate peroxidase, superoxide dismutase (SOD), and catalase (CAT) (Chen and Gallie, 2004; Desikan *et al.*, 2007; Jahan *et al.*, 2008; Miao *et al.*, 2006; Neill *et al.*, 2002). Glutathione, ascorbate and their corresponding peroxidases reduce hydrogen peroxide in the ascorbate-glutathione and glutathione-peroxidase cycles, while CAT reduces H_2O_2 producing water and molecular oxygen (Apel and Hirt, 2004). The role of NO as an antioxidant in plants is less clear: on one hand (as mentioned above) NO suppresses the activity of catalase and ascorbate peroxidase, but on the other hand NO can enhance the activity of SOD and itself can react with superoxide to form peroxynitrite (Neill

et al., 2008), so it would seem that NO acts against superoxide and in favour of hydrogen peroxide. Other methods of ROS control may include curbing its production by NADPH-oxidase dephosphorylation or allosteric inactivation (eg nitrosylation) (Belin *et al.*, 2006; Gosti *et al.*, 1999; Merlot *et al.*, 2001; Yoshida *et al.*, 2006; Yun *et al.*, 2011).

ROS function as signalling molecules in response to a variety of stimuli such as ABA, ethylene, salicylic acid (SA), ozone, or jasmonic acid (JA) (Taiz and Zeiger, 2002). Hydrogen peroxide is the most common redox signalling molecule because it is moderately reactive (the other ROS are far more reactive and not ideal for signalling), has a relative long half-life (1ms), it can diffuse across cell membranes, its production can be induced and stopped quickly, and can cause the post-translational modification of proteins to alter their activity (Marks *et al.*, 2009; Wang and Song, 2008). In addition to baseline metabolic production of ROS, cells have dismutases and NADPH-oxidases to produce ROS specifically for signalling (Kwak *et al.*, 2003). Furthermore, in specialised cells such as guard cells with negligible photosynthetic activity, the chloroplasts can be used as a source of ROS for signalling (Pham and Desikan, 2009). ROS have been found to alter the activity of two-component systems in a thiol group-dependent way, which may result in expression or repression of target genes (Apel and Hirt, 2004; Desikan *et al.*, 2006). Hydrogen peroxide can also activate several MAPK-cascades (for which there are over 100 genes in *Arabidopsis*), regulate the activity of ion channels and proton pumps which control pH and membrane polarity (Apel and Hirt, 2004; Cho *et al.*, 2009b; Pham and Desikan, 2009). The sheer variety of targets for interaction makes ROS a vital component of many signal transduction pathways.

2.2.4 Nitric oxide

Nitric oxide (NO) is a reactive gas molecule and a free-radical (sometimes represented as $NO^{\bullet}$) present in many cells and tissues. In most cells NO is produced from L-arginine by the nitric oxide synthase (NOS) family of enzymes. Although in plants NOS-like enzymes have not been found, NO production in plants has been shown to occur via nitrate reductase (NR) enzymes such as NIA1 and NIA2 (Desikan *et al.*, 2002), and others such as AtNOA1, and nitrite reductase (Wilson *et al.*, 2008).

As with ROS, NO has the attributes of a second messenger: it is moderately reactive, has a half life of 5-10s, can diffuse across membranes and in lipid environments, and its concentration can be controlled by the cell either by stimulating its production (eg with NOS or NR) or by accelerating its removal (Hancock, 1999; Wilson *et al.*, 2008).

As a signalling molecule NO acts on guanyl-cyclase by binding on haem groups to produce cyclic guanosine monophosphate (cGMP), another second messenger that regulates serine threonine kinases. NO can induce conformational change in proteins by nitrosylation of cystenyl-SH groups which can activate or deactivate an enzyme (a conformational change analogous to phosphorylation), control ion channels, elevate cytosolic Ca^{2+}, and interact with ROS (Marks *et al.*, 2009; Wilson *et al.*, 2008). The effect of NO depends critically on its concentration. For example, in guard cells NO can deactivate inwards K^+ channels, but in higher doses it can deactivate outwards K^+ channels as well (Bright, 2006; Sokolovski and Blatt, 2004). A halt of ion efflux can possibly block stomatal movements. When NO is overproduced nitrosylation can occur non-specifically (a situation known as nitrosative stress, similar to oxidative stress), and can trigger cell death; NO induced cell death is used as a defence mechanism against invading organisms (Marks *et al.*, 2009). In seeds, NO antagonises ABA by reducing seed dormancy and stimulating germination (Bethke *et al.*, 2006; Libourel *et al.*, 2006).

NO can be removed from cells by combining it with water, oxygen, and haemoglobins to produce nitrites and nitrates. The antioxidant glutathione can bind to NO to from S-nitrosylated glutathione (GSNO) which is converted to ammonia by GSNO-reductase. In addition to reacting with an antioxidant, NO itself can have antioxidant activity by reacting with superoxide to create peroxynitrite, normally a toxic substance but surprisingly plant cells can resist its toxicity and even use it as a messenger (Delledonne *et al.*, 2001; Klotz *et al.*, 2002).

2.2.5 Calcium

Calcium ions (Ca^{2+}) have important roles in signalling processes in all cells (Hancock, 1999). The cytosolic concentration of Ca^{2+} is kept low in unstimulated conditions by means of calcium pumps, channels, stores, and buffering proteins. ATP-dependent pumps

such as the Sarco/Endoplasmic Reticulum Calcium-dependent ATPase (SERCA) and the plasma membrane Ca^{2+}-dependent ATPase (PMCA) transport Ca^{2+} ions against their gradient into the Endoplasmic Reticulum (ER) and the outside of the cell, respectively. Na^{+}–Ca^{2+} and H^{+}–Ca^{2+} exchangers use the gradient in the concentration of Na^{+} and H^{+} to transport Ca^{2+} either outside the mitochondria or outside the cell. Ca^{2+}-uptake into the mitochondria occurs via a uniporter (UP) along a charge gradient when required by a cellular process. Ligand-controlled channels can release calcium ions into the cytosol from the ER and the exterior of the cell. One of the best-known ligands controlling Ca^{2+} from the ER is inositol-3-phosphate (IP3). Although IP3 is active in guard cell signalling its receptors in plant genomes have yet to be found (Kim *et al.*, 2010). Additionally, voltage-controlled channels can import Ca^{2+}-ions from the exterior into the cytosol (Marks *et al.*, 2009). In plant cells, vacuoles also serve as calcium stores (Taiz and Zeiger, 2002). In addition to stores and channels, cells also control cytosolic Ca^{2+} concentration via buffering proteins such as calmodulin. All the above mechanisms assure that under unstimulated conditions, the concentration of calcium ions in the cytosol is four orders of magnitude lower than outside the cell or in the stores, 10^{-7} M in the cytosol versus 2×10^{-3} M outside the cell and in the stores (Hancock, 1999).

Many Ca^{2+}-binding proteins contain a motif known as the EF-hand, consisting of a helix-loop-helix domain where calcium ions can bind. In the *Arabidopsis* genome there are over 200 Ca^{2+}-binding proteins, including 34 calcium-dependent protein kinases (CDPK) and 10 calcineurin B-like (CBL) proteins (D'Angelo *et al.*, 2006; Kim *et al.*, 2010). Interestingly, in *Arabidopsis* NADPH-oxidases have an EF-hand which is activated by cytosolic Ca^{2+} (Ogasawara *et al.*, 2008). Many proteins upon binding Ca^{2+} undergo conformational changes which might expose (or hide) active sites, altering the activity of the protein and downstream signalling components.

Calcium signalling does not exclusively depend on cytosolic concentration. Many signals induce transients, oscillations, and location-specific changes that have different effects on cell behaviour, making Ca^{2+}-signalling an extremely versatile tool for signalling. It has been hypothesised that some signals induce specific "calcium signatures" to produce specific cellular responses (Ng *et al.*, 2001; Taiz and Zeiger, 2002). Furthermore, it has been shown that unstimulated *Arabidopsis* guard cells exhibit cytosolic

Ca^{2+}-oscillations (Yang *et al.*, 2008), giving rise to the hypothesis that calcium signalling may depend less than previously thought on concentration and may involve the "priming" of Ca^{2+} sensors (Kim *et al.*, 2010; Young *et al.*, 2006).

2.2.6 Ions

Ions are particles with an imbalance in their numbers of electrons and protons which result in an electric charge. If there are more protons than electrons then the charge is positive, if the reverse is true then the charge is negative. There are five main ions in cells: H^+, Na^+, K^+, Cl^-, and Ca^{2+} (Marks *et al.*, 2009). Ions are used by cells to control pH (H^+), membrane potential (Na^+, K^+, Cl^-, Ca^{2+}), and function as second messengers (Ca^{2+}). Additionally, the concentration of K^+ and Cl^- in guard cells is regulated to establish osmotic gradients to change the amount of water contained in the vacuoles which ultimately determines cell volume and stomatal aperture (Taiz and Zeiger, 2002).

The cytosolic concentration of K^+ ions in guard cells is controlled by the inwards-rectifying K^+_{in} and the outwards-rectifying K^+_{out} channels, located in the plasma membrane (Blatt and Armstrong, 1993; Schroeder and Fang, 1991), and slow (SV) and voltage independent (VK) K^+ vacuolar channels (MacRobbie, 1998; Ward and Schroeder, 1994). The activity of these channels is controlled by pH, membrane potential (and vacuolar membrane potential), and cytosolic Ca^{2+} concentration.

The channels regulating the concentration of Cl^- anions are located on the plasma membrane and include the slow-activating (S-type) and the rapid transient (R-type) (Hedrich *et al.*, 1990; Schroeder and Hagiwara, 1989). Vacuolar anion channels controlled by CDPKs are reportedly active during stomatal closure (MacRobbie, 1998; Pei *et al.*, 1996).

The concentration of H^+ (and hence pH) in the cytosol is mainly controlled by proton pumps located in the plasma membrane (H^+–ATPase) and in the vacuolar membrane (V–ATPase). Variations in the pH of the cytosol has consequences for the behaviour of the cell and the activity of the proteins, eg more alkaline pH increases the availability of K^+ outwards channels (see below) (Blatt and Armstrong, 1993).

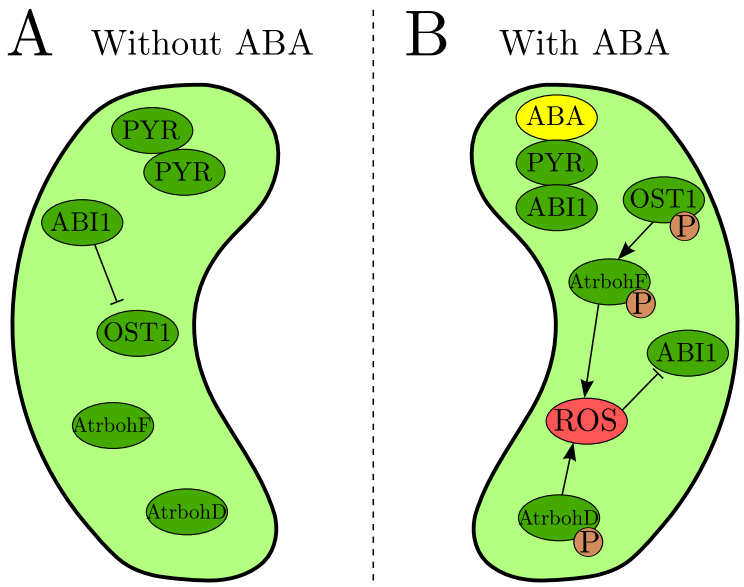

Figure 2.3: Early events of ABA signalling in guard cells. **A:** In the absence of ABA, PYR molecules form dimers and do not interact with ABI1, which dephosphorylates the kinase OST1. The NADPH-oxidase AtrbohF does not produce ROS because OST1 activity is blocked by ABI1. AtrbohD does not produce ROS in the absence of ABA. **B:** The ABA-bound PYR/PYL impede ABI1 inactivation of OST1, which phosphorylates and induces ROS production by AtrbohF. ABA also induces AtrbohD-mediated production of ROS.

2.3 Abscisic acid-induced stomatal closure

2.3.1 ABA perception

The family of proteins pyrabactin resistance (PYR) and PYR-like (PYL), sometimes also known as regulatory component of ABA receptor (RCAR) have been shown to be receptors of ABA in guard cells (Ma *et al.*, 2009; Park *et al.*, 2009). The perception of ABA and early events in ABA signalling form a module that includes both positive and negative regulators of the pathway (Hubbard *et al.*, 2010; Weiner *et al.*, 2010). The protein phosphatase 2C (PP2C) ABA insensitive 1 (ABI1) is known to be a negative regulator of ABA signalling, and lies upstream of all other known signalling events and components involved in stomatal closure (Allen *et al.*, 1999; Gosti *et al.*, 1999; Merlot *et al.*, 2001; Pei *et al.*, 1997). ABI1 inactivates the kinase activity of the SNF1-related protein kinase 2 (SnRK2) open stomata 1 (OST1) by removing phosphate groups (Mustilli *et al.*, 2002;

Vlad *et al.*, 2009; Yoshida *et al.*, 2006). In the absence of ABA, proteins of the PYR/PYL family form dimers (Santiago *et al.*, 2009), and ABI1 prevents OST1 from activating other components that lie downstream in the pathway (Fig. 2.3A). The structure of PYR/PYL proteins contains a cavity where the ABA binding site is located. Upon arrival of the ABA signal, ABA molecules bind to PYR/PYL proteins in the cavity which weakens and possibly breaks the dimer; ABA molecules are secured to the protein by a gate and lock mechanism (Melcher *et al.*, 2009, 2010; Miyazono *et al.*, 2009; Nishimura *et al.*, 2009; Santiago *et al.*, 2009). Binding of ABA, in addition to weakening the dimer, also changes the conformation of PYR/PYL molecules which increases their affinity to PP2C molecules (Miyazono *et al.*, 2009). ABA-bound PYR/PYL have high affinity to ABI1, and inhibit its phosphatase activity, thus allowing OST1 to become phosphorylated and activate downstream components (Fig. 2.3B) (Weiner *et al.*, 2010). PYR/PYL molecules are small and soluble (Hubbard *et al.*, 2010), which indicates that ABA must enter the cytosol to bind to its receptors. A mediator of ABA uptake into the cell is the ATP binding cassette (ABC) sometimes called *Arabidopsis thaliana* pleiotropic drug resistance transporter (PDR12) located on the plasma membrane (Kang *et al.*, 2010).

The Mg-Chelatase H-subunit (CHLH), located in the chloroplasts has been identified by Shen *et al.* (2006) as an ABA receptor in *Arabidopsis* guard cells; however, this result has been contested in barley (Müller and Hansson, 2009). Additionally two G-protein coupled receptor type G-proteins GTG1 and GTG2, located on the membrane, have been identified as ABA receptors (Pandey *et al.*, 2009). Although GTG1 and GTG2 are involved in stomatal closure, loss-of-function *gtg1 gtg2* single and double mutants still retain some ABA sensitivity (Pandey *et al.*, 2009), possibly via PYR/PYL proteins. The precise interaction between GTG1, GTG2, and the rest of the guard cell signalling pathway remains to be uncovered.

It is perhaps worth mentioning that the search for the ABA receptors has not been straightforward (McCourt and Creelman, 2008). For example, the RNA-binding protein FCA was identified as an ABA receptor (Razem *et al.*, 2006), which later was shown to be incorrect (Razem *et al.*, 2008; Risk *et al.*, 2008). The role of CHLH needs also to be clarified, given the negative result from the study in barley. So far, only the role of PYR/PYL, via their interactions with ABI1 and OST1, is firmly established in the

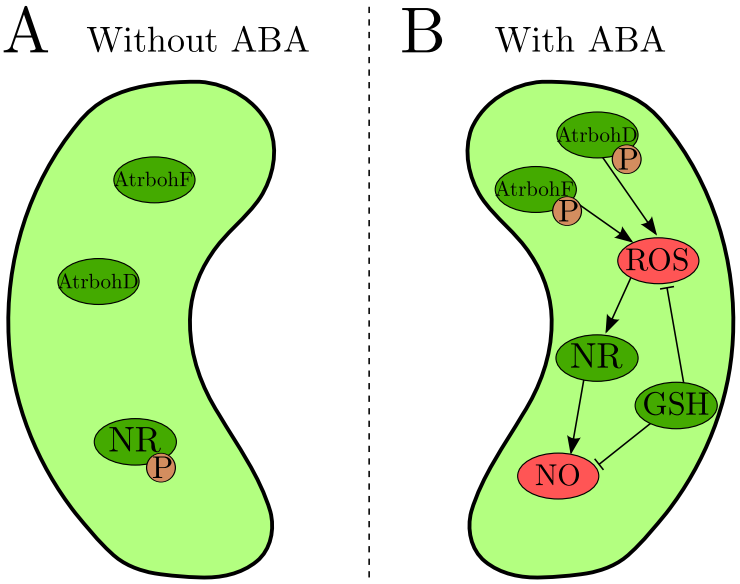

Figure 2.4: ABA-induced ROS and NO synthesis in guard cells. **A**: In the absence of ABA AtrbohD/F remain inactive so ROS are not produced. Without ABA-induced ROS NR remains in a phosphorylated state and there is no NO produced for signalling. **B**: When the ABA signal arrives AtrbohD/F phosphorylate and ROS is produced. ROS cause the dephosphorylation of NR which produces NO. Gluthathione (GSH) scavenges both ROS and NO.

ABA-induced stomatal closure signalling pathway.

2.3.2 ROS signalling in ABA-induced stomatal closure

ABI1 impedes the activation of OST1 in the absence of an ABA signal (Fig. 2.3A). After ABA perception by guard cells and the phosphatase activity of ABI1 has been halted, OST1 is able to autophosphorylate and become active. Although in *Arabidopsis* there are at least ten members of the NADPH-oxidases *Arabidopsis thaliana* respiratory burst oxidase homolog (Atrboh) only two versions have been found to be active in ABA-induced ROS production: AtrbohF and AtrbohD. It is known that AtrbohF is phosphorylated by OST1 to produce superoxide which quickly becomes dismutated into hydrogen peroxide (Fig. 2.3B) (Belin *et al.*, 2006; Kwak *et al.*, 2003; Mustilli *et al.*, 2002). It is also known that AtrbohD is active in ABA-induced stomatal closure, although its activation mechanism is still unknown. Studies of loss-of-function mutants *atrbohd* and *atrbohf*, have shown that

the amount of ROS produced by AtrbohD is less than the amount produced by AtrbohF, and *atrbohd* mutants have sensitivity to ABA comparable to wild-type specimens (Kwak *et al.*, 2003). ROS have also been reported to inhibit the phosphatase activity of ABI1 *in vitro*, possibly creating a positive feedback loop (Meinhard and Grill, 2001), whose existence needs to be confirmed *in vivo*. Hydrogen peroxide production in guard cells begins within a minute of exposure to ABA in an oxidative burst (Pei and Kuchitsu, 2005; Pei *et al.*, 2000; Wang and Song, 2008). The release of Ca^{2+} and its effect on ion channels (see below), and subsequent stomatal closure in response to ABA depends on ROS, and ROS-regulated MAPK proteins (Jammes *et al.*, 2009; Köhler *et al.*, 2003; Pei *et al.*, 2000). Other targets of ROS include glutathione peroxidase (GPX), which acts as a transducer and scavenger of ROS (Miao *et al.*, 2006), and possibly cell pH.

Antioxidants ascorbate (ASC) and glutathione (GSH) have been found in guard cells, and are involved in the prevention of oxidative stress damage that might arise from excessive ROS accumulation in guard cells (Fig. 2.4B) (Chen and Gallie, 2004; Kwak *et al.*, 2006). The expression of dehydroascorbate reductase (DHAR) which produces ASC, for example, is regulated by circadian rhythms to allow ROS concentrations to increase in the evening to allow stomatal closure at night (Chen and Gallie, 2004). As mentioned above, GPX has a dual role in ROS signalling, its action is represented by Miao *et al.* (2006) as a balance scale: when there is little ROS then GPX acts as a transducer, once the ROS concentration passes a tipping point then GPX becomes a scavenger. The exact way in which these antioxidants interact with the rest of the components of the ABA signalling pathway remains to be clarified. ROS are now firmly established as an essential component of the ABA-induced stomatal closure, but their regulation and removal during stomatal closure needs to be understood yet (Kwak *et al.*, 2006; Pham and Desikan, 2009; Wang and Song, 2008; Zhang *et al.*, 2001b).

2.3.3 NO production

Nitric oxide signalling occurs during biotic and abiotic stress responses in plants, and ABA-induced stomatal closure is not an exception (García-Mata and Lamattina, 2002; Melotto *et al.*, 2006; Neill *et al.*, 2003). Production of NO in guard cells is observed after treatment

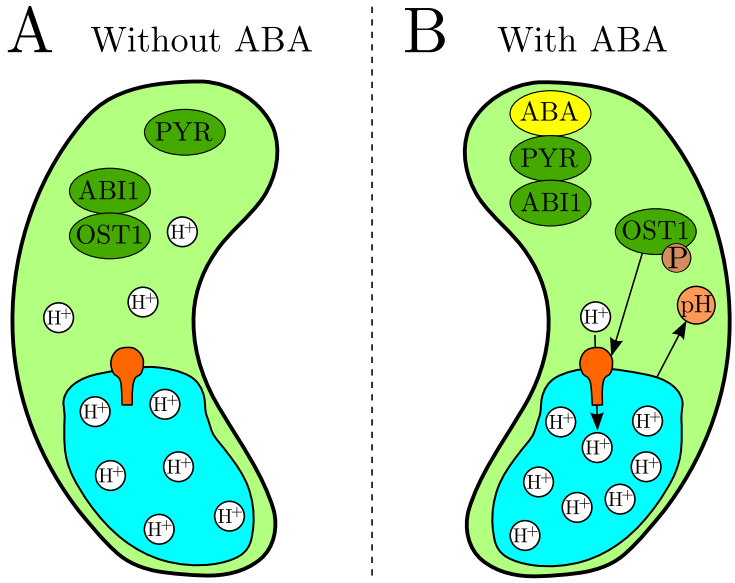

Figure 2.5: ABA-induced cytosolic alkalinisation. **A**: In the absence of ABA the phosphatase ABI1 inactivates OST1, and cytosolic and vacuolar pH are kept at equilibrium. **B**: ABA-bound PYR/PYL blocks OST1 inactivation of ABI1. Phosphorylated OST1 increases the activity of the vacuolar proton pump, making the pH of the cytosol more alkaline and the pH of the vacuole more acidic.

with ABA, and cells exposed to NO-scavengers become less responsive to stimuli. Loss-of-function single and double NR mutants *nia1nia2* fail to produce NO and show decreased stomatal responses to ABA stimuli (Desikan *et al.*, 2002), and production of NO by NR *in vitro* has been shown to be modulated by phosphorylation (Rockel *et al.*, 2002). In ABA-induced stomatal closure NO production depends on hydrogen peroxide generated by AtrbohD/F (Bright *et al.*, 2006) (Fig. 2.4). ROS-dependent NO production is intriguing given that, as explained above, NO and superoxide react with each other. This interaction suggests that ROS may modulate their own concentration through NO. Additionally, NO and hydrogen peroxide share GSH as a scavenger (Perazzolli *et al.*, 2006; Wilson *et al.*, 2008). ABA-induced NO in guard cells regulates the concentration of K^+ and Cl^- by facilitating the release of Ca^{2+} from intracellular stores (Fig. 2.6, see below) (García-Mata *et al.*, 2003).

2.3.4 Cytosolic alkalinisation

During ABA-induced stomatal closure the pH of the cytosol becomes more alkaline, going from around 7.0 to 7.5; whereas the pH in the tonoplast (vacuoles) becomes more acidic, decreasing from 5.5 to 5.0 (Blatt and Armstrong, 1993; Islam *et al.*, 2010; Miedema and Assmann, 1996; Pei *et al.*, 1998; Zhang *et al.*, 2001a). Changes to cytosolic pH during stomatal closure are driven mainly by the vacuolar proton pump (V–ATPase), given that rises in Ca^{2+} deactivate the plasma membrane H^{+}–ATPases (Sec. 2.3.5). According to what is described in Sec. 2.3.1, OST1 is inactivated by ABI1 in the absence of ABA (Fig. 2.5A). When ABA-bound PYR/PYL prevents OST1 dephosphorylation, then OST1 is able to increase V–ATPase activity (Fig. 2.5B), though the details of this interaction remain to be established (Islam *et al.*, 2010; Suhita *et al.*, 2004). There are conflicting reports in the literature about the involvement of ROS in the rise of cytosolic pH after treatment with ABA, which suggests complex interactions of the signalling components at this level: Suhita *et al.* (2004) report that alkalinisation of the cytosol precedes ROS production, whereas Zhang *et al.* (2001a) find that to the contrary, pH rises in response to ROS elevations. What is not disputed is that cytosolic pH has an important role in ABA-induced stomatal closure, and the effects of it becoming more alkaline include the regulation on ion channels (see below).

2.3.5 Ca^{2+} release

The role of Ca^{2+} in ABA-induced stomatal closure was established by McAinsh (1990) in a study where the cytosolic calcium levels were observed to rise in 80% of cells treated with ABA. ABA-induced ROS production activates plasma-membrane Ca^{2+} channels (Fig. 2.6) (McAinsh *et al.*, 1996; Murata *et al.*, 2001; Pei *et al.*, 2000). In addition, OST1 is thought to interact with Ca^{2+} channels, affecting cytosolic Ca^{2+} levels in a ROS-independent way (Kim *et al.*, 2010; Siegel *et al.*, 2009). NO also mobilises Ca^{2+} into the cytosol in a process that requires protein phosphorylation (García-Mata *et al.*, 2003; Lamotte *et al.*, 2005; Neill *et al.*, 2008; Sokolovski *et al.*, 2005). As mentioned in Sec. 2.2.5, unstimulated cells have spontaneous cytosolic Ca^{2+} oscillations, and ABA-induced calcium increases also feature oscillations (Kim *et al.*, 2010). Furthermore, it has

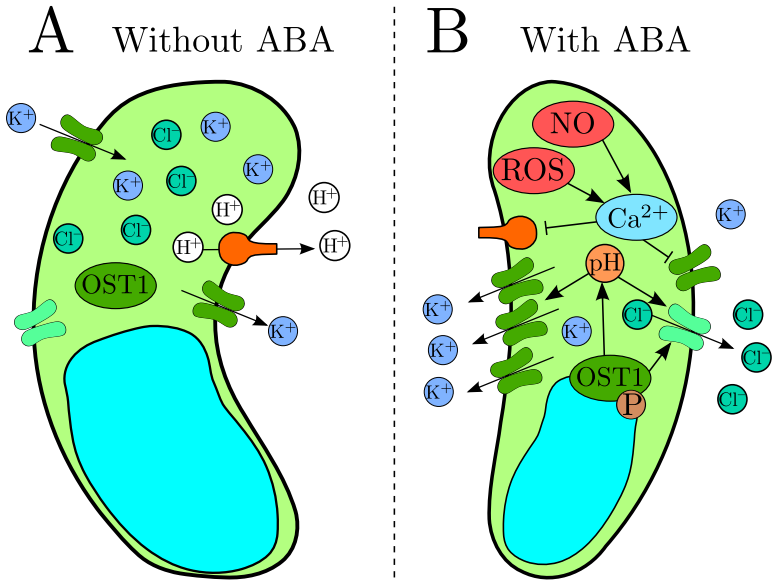

Figure 2.6: ABA-induced ion efflux and loss of turgor. **A:** In the absence of ABA the kinase OST1 is inactive and ion levels and pH remain at equilibrium levels. **B**: ABA-induced ROS and NO mediate the rise in cytosolic Ca^{2+}, membrane depolarisation, and ion efflux. Active OST1 causes pH to rise and activates anion channels. The resulting osmotic gradient causes the exit of water from the cell and the stomatal pore to close.

been shown that elevations in the cytosolic level of calcium are not required to successfully complete ABA-induced stomatal closure, though the presence of calcium at least at resting levels is required (Siegel *et al.*, 2009). These observations have prompted the hypothesis that ABA "primes" the Ca^{2+} receptors to be more sensitive (Kim *et al.*, 2010). Downstream in the ABA signalling pathway Ca^{2+} controls ion channels (Ng *et al.*, 2001) and membrane polarity (see below). Observations by Levchenko *et al.* (2005), McAinsh (1990) and Siegel *et al.* (2009) also show that there are Ca^{2+}-independent branches in the guard cell signalling network, eg via pH.

2.3.6 Membrane polarity, ion and anion efflux

ABA-induced rises in the levels of ROS and NO promote the influx of Ca^{2+} into the cytosol through membrane channels and its release from intracellular stores. This calcium elevation down-regulates the activity of inwards-rectifying channels (K^+_{in}) (Fig. 2.6) (García-Mata *et al.*, 2003; Kwak *et al.*, 2008; Pei *et al.*, 2000; Schroeder and Hagiwara, 1989). A

rise in cytosolic pH and an elevated NO concentration cause a reduction in the cytosolic concentration of K^+ by increasing the efflux through K^+_{out} and a reduced influx through K^+_{in}. More alkaline levels of pH increase the number of available K^+_{out} channels to accelerate the rate of K^+ ion extrusion (Blatt and Armstrong, 1993; Miedema and Assmann, 1996). Very high concentrations of NO ($\geq$ 100nM), however, can block K^+_{out} directly by nitrosylation (García-Mata *et al.*, 2003; Sokolovski and Blatt, 2004; Sokolovski *et al.*, 2005), possibly compromising the closure process. High ROS concentrations can block K^+_{in} (Zhang *et al.*, 2001c) and, surprisingly, K^+_{out}. However, it has been suggested that the blocking of K^+_{out} by ROS may not be a feature in ABA-mediated stomatal closure (Köhler *et al.*, 2003). Anions such as Cl^- and malate are released from the vacuole and out of the cell during stomatal closure (Assmann and Shimazaki, 1999; Wang *et al.*, 2001). OST1 phosphorylates the slow anion channel-associated 1 (SLAC1) channel, causing an Cl^- efflux which depolarises the cell membrane (Geiger *et al.*, 2009). Guard cells dispose of malate during stomatal closure by extrusion and consumption by the tricarboxylic acid cycle (Dittrich and Raschke, 1977). An additional consequence of the rise in cytosolic Ca^{2+}, NO, and ROS is the deactivation of membrane proton pumps (H^+-ATPases), which further contribute to depolarise the membrane (Wasilewska *et al.*, 2008), and allows K^+_{out} to increase its activity.

2.3.7 Loss of turgor and stomatal closure

The ion and solute efflux described above, and the resulting lower ionic concentration in the cytosol create an osmotic gradient that forces the water in the vacuoles to exit the cell. The loss of water leads to a decreased turgor pressure thus deflating the guard cells and closing the pore (Kwak *et al.*, 2008; Schroeder *et al.*, 2001). The loss of turgor marks the culmination of the stomatal closure signalling process and, depending on circumstances, may mark the beginning of inhibition of stomatal opening, a related but distinct process from stomatal closure (Kim *et al.*, 2010; Pandey *et al.*, 2009).

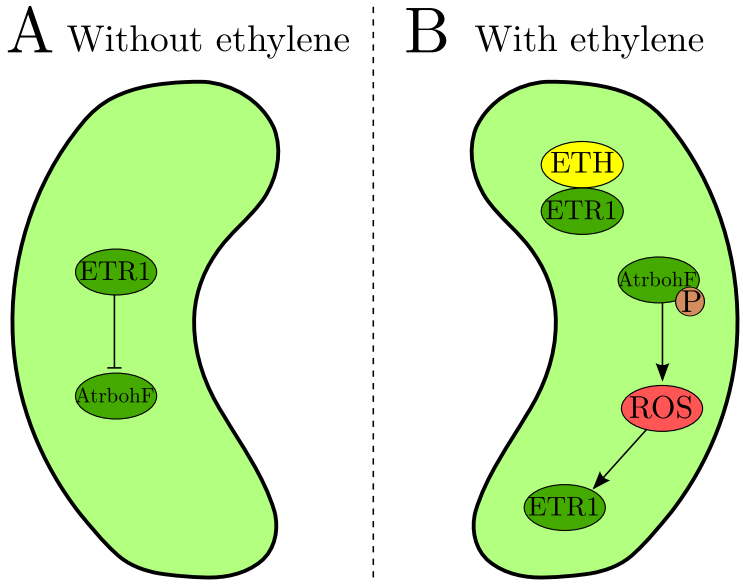

Figure 2.7: Ethylene reception and ROS production in guard cells. **A**: ETR1 is a negative regulator of ethylene signalling. **B**: Ethylene binding inactivates ETR1, ROS is produced by AtrbohF. ETR1 senses ROS and relays the signal downstream.

2.4 Ethylene-induced stomatal closure

2.4.1 Ethylene

As described in Sec. 2.2.2, ethylene is an important and well-studied plant hormone. In addition to all the functions already mentioned, ethylene is an effector of stomatal closure in *Arabidopsis thaliana* (Pallas and Kays, 1982). Intriguingly, in other plant species such as *Vicia faba* ethylene is an effector of stomatal opening as mentioned in Sec. 2.2.2, however this work is focused only in ethylene effects in *Arabidopsis*. Although the ethylene pathway is arguably well-understood in seedlings and root cells (Stepanova and Alonso, 2005), little is known about it in guard cells. Below is described what is known so far about the ethylene signalling pathway in guard cells leading to stomatal closure.

2.4.2 Ethylene perception and ROS production

Of the five known ethylene receptors only ETR1, located in the ER and a negative regulator of ethylene signalling, has been shown to be involved in ethylene-induced

stomatal closure (Desikan *et al.*, 2005; Tanaka *et al.*, 2005, 2006). The immediate events after ethylene binding by ETR1 are still not known. Ethylene induces ROS production by AtrbohF but, unlike in the ABA-pathway, not by AtrbohD (Fig. 2.7) (Desikan *et al.*, 2006). The signalling steps between ethylene perception by ETR1 and ROS production by AtrbohF are not yet clear. Given that ETR1 is a negative regulator of the ethylene pathway (ie ethylene inactivates ETR1), it is probable that its action on AtrbohF (in the absence of ethylene, Fig. 2.7A) is also negative. Enzymes such as ABI1 and OST1 could be involved in the production of ROS by ethylene (which would make the earliest events in ethylene signalling resemble the early events in ABA signalling), although that remains to be confirmed. In addition to ethylene perception, ETR1 has an additional role in ethylene signalling: it senses ROS and relays the signal downstream in a process that depends on the enzymes ethylene insensitive 2 (EIN2) and *Arabidopsis thaliana* response regulator 2 (ARR2) (Fig. 2.8); however, their exact roles have not been determined so far (Desikan *et al.*, 2006).

2.4.3 Downstream of ROS

The events downstream of ethylene-induced ROS production during stomatal closure have not been fully verified experimentally. Preliminary results indicate that NO is produced in guard cells treated with the ethylene precursor ACC (Devienne, 2010), possibly in a pH-dependent way (Liu *et al.*, 2010a). If indeed NO is produced, it is likely that at least some is produced by NIA1 via ROS as it occurs in the ABA pathway, but other sources may also be involved. The role of Ca^{2+}, membrane polarity, and ion channels in ethylene-induced stomatal closure has not been experimentally tested. An osmotic gradient must be produced so that water may exit the cells for stomatal closure to happen, this requires ion extrusion (Fig. 2.8). Given the overlap of the ethylene and ABA pathways at the AtrbohF and ROS level, it is likely that the components downstream of ABA are also downstream of ethylene.

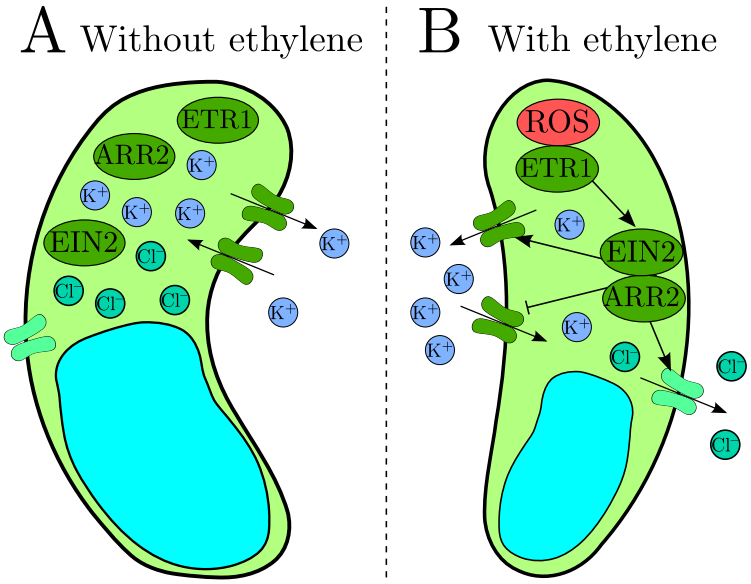

Figure 2.8: Ethylene signalling downstream of ROS. **A**: At rest the cells maintains equilibrium levels of ROS, and ion concentration. **B**: Ethylene-induced ROS requires ETR1, EIN2, and ARR2 so that stomatal closure can be completed, which implies that the extrusion of ions from the cytosol and other cellular compartments must occur.

2.5 Lack of closure under combined ABA-ethylene treatment

Hormone interactions in plant cells are common (Acharya and Assmann, 2009; Chang and Bleecker, 2004; Liu *et al.*, 2010b; Wilkinson and Davies, 2009a); Sections 2.3 and 2.4 discuss the mode of action of ABA and ethylene in guard cells. Both of these hormones cause stomatal closure individually but, surprisingly, when applied simultaneously they fail to produce full closure (Desikan *et al.*, 2006; Tanaka *et al.*, 2005). These observations are hard to reconcile; two signals that individually produce a particular response but none when they are combined seems contrary to intuition, specially since their signalling pathways overlap. The mechanisms by which this cross-talk occurs are not yet known. The signal transduction pathways presented in this chapter are only a partial description of stomatal closure whose components are thought to be the most important ones *in vivo*. An exhaustive list of all known interactions, whose understanding is incomplete in many cases and sometimes fuzzy at best, is not presented here. For example, the roles of ARR, EIN2, and ETR1 in the ABA pathway so far have not been elucidated. In addition, there are components of the ethylene pathway that have been known to interact with or

antagonise ABA signals in other cell types, for example: EIN2 modulates ABA stress responses in *Arabidopsis* seeds (Wang *et al.*, 2007), ABA can restrict ACC production also in seeds (Matilla, 2000; Yoshii and Imaseki, 1981), and oxidative stress (ie excess ROS) and ethylene are suspected to suppress stomatal closure (Wilkinson and Davies, 2009b). The picture of stomatal closure is much more complicated than what is understood today.

Large and complex signalling systems which are almost always non-linear are known to produce strange and unexpected behaviour such as bistability, oscillations, and chaos (Carter *et al.*, 2010; Stark and Hardy, 2003; Stark *et al.*, 2007; Strogatz, 1994). In a way, it is not surprising to find seemingly contradictory behaviour in a network with as many nodes and interactions as the guard cell network.

In the rest of this thesis we will endeavour to understand the possible causes of the ABA-ethylene cross-talk in guard cells. Chapter 3 provides a minimal introduction to mathematical modelling for cell-signalling systems. Chapter 4 introduces a new method to fit model parameters from data using Monte Carlo techniques and evolutionary algorithms. Models of activation cascades are analysed in Chapter 5; analytical solutions are obtained, and their uses in different contexts are discussed. Experimental measurements are made in guard cells treated with single and combined stimuli in Chapter 6; an ordinary differential equation model of stomatal closure is constructed in Chapter 7, incorporating the results from the analysis of activation cascades and fit to the data using the new Monte Carlo method.

Chapter 3

Mathematical models in cell biology

Systems studied in the biological, physical, and social sciences usually contain many components that interact in specific, yet complicated ways, often changing with time. In biology, cell signalling systems (eg the regulation of transpiration in plants described in Chapter 2 or the immune system) have numerous components including nucleic acids, proteins, hormones, reactive molecules and ions (Hancock, 1999). These components interact to elicit responses to external signals. For example, in sections 2.3 and 2.4 we discussed how when plants detect unfavourable conditions (eg drought or pathogen attack), interactions between various cell signalling components act to slow down their metabolism until conditions improve. Although many components of these systems are known, the size and complexity of the system often make detailed experimental investigations too expensive or difficult to perform (Hardy and Stark, 2002; Stark and Hardy, 2003; Yates *et al.*, 2001). The use of tools from quantitative and physical disciplines has helped to overcome some of these limitations by abstracting and representing biological systems in a framework that allows the efficient exploration of hypotheses, and the extraction of conclusions that would be difficult to obtain by experimentation alone. In biology, the need for a quantitative understanding is specially acute. Mathematical modelling has been successfully used to explain and predict system behaviour, what experiments should be performed, what are the 'rate-limiting' components of a system and whether or how a system can be controlled (Howison, 2003; Keener and Sneyd, 1998). In many systems, such as guard cells, many questions remain without an answer, such as what is the precise

timing of events that lead to a response? Which components form groups of interaction, when and for how long? What determines specificity to multiple stimuli?

Broadly, a model is an abstract representation of a phenomenon using in this case a mathematical framework. There is no unique way to model a system, and rarely a model is *the model* that answers all questions, nor is there a recipe for producing mathematical models; generally, the process involves making assumptions about what the most important characteristics of a system are, and ignoring the other less-important aspects of it. A good model will capture the essential characteristics of the system, make predictions, and guide new experiments. Models can always be refined further as needed, but should always be kept as simple as possible.

The type of models that we use in this work are mechanistic and deterministic; they describe a mechanism for the behaviour of a system. Once the rules of the system are set and we provide some necessary initial or boundary conditions, the behaviour of the system is determined for all time. Examples of deterministic models include differential equations (continuous time), and Boolean models and difference equations (discrete time). Other type of approaches are mechanistic but incorporate randomness, such as stochastic models. Statistical models describe observations and data but do not provide insights into the mechanisms that drive systems.

3.1 Differential equation models

In our modelling, the state of a biological system is described by an n-dimensional continuous function $\mathbf{x}(t) = [x_1(t), x_2(t), \ldots, x_n(t)]$, where t is time, and the variables $x_i(t)$ are features of the system such as temperature or concentration of a substance. We represent the change of $\mathbf{x}(t)$ in time using *ordinary differential equations* (ODEs), which describe the rate of change of an unknown function along one dimension, usually time (Kreyszig, 2006). *Partial differential equations* (PDEs) describe the change along more than one dimensions (eg time and spatial dimensions). In this project we are concerned mostly with the change of a function in time, hence we will use ODEs. We write the change of $\mathbf{x}$ in time as

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = f(t, \mathbf{x}, \boldsymbol{\theta}), \tag{3.1}$$

where $\frac{\mathrm{d}}{\mathrm{d}t}$ is the time-derivative of $\mathbf{x}$. We relate the rate of change of $\mathbf{x}$ in time to a function of time t, its current state $\mathbf{x}(t)$, and a set of parameters $\boldsymbol{\theta}$. Parameters are constants that stand for the inherent properties of the system, and typically are non-negative real numbers (the signs are prescribed by the equations). To solve equation (3.1), we need to know the *initial conditions*, the state of $\mathbf{x}$ at some time t_0, ie $\mathbf{x}(t_0) = \mathbf{x}_0$.
Equation (3.1) is *linear* whenever f is a linear operator on $\mathbf{x}$, ie when the following properties hold

$$f(t, \mathbf{x} + \mathbf{y}, \boldsymbol{\theta}) = f(t, \mathbf{x}, \boldsymbol{\theta}) + f(t, \mathbf{y}, \boldsymbol{\theta}),$$
$$f(t, \alpha\mathbf{x}, \boldsymbol{\theta}) = \alpha f(t, \mathbf{x}, \boldsymbol{\theta}),$$

and α is a fixed constant. If f does not have these properties then the equation is *nonlinear*. Most biological systems present nonlinear behaviour; however, there are several cases in which a linear model or a linear approximation can be used to understand the system. Equation (3.1) reaches a steady state or an equilibrium point at $\mathbf{x}_{ss}$ when:

$$f(t, \mathbf{x}_{ss}, \boldsymbol{\theta}) = 0,$$

that is, the rate of change is zero and the state of the system does not change. A steady state can be stable, ie if the system is perturbed from the steady state it will return to it, or unstable, ie if system does not return to the steady state when perturbed. Other ODE systems may exhibit other more complex behaviour such as regular and irregular oscillations (Strogatz, 1994).

Below, we present an example that is simple enough so that $x(t)$ can be written explicitly.

Example 3.1.1. Suppose a one-dimensional system $x(t)$ whose rate of change is given by

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \alpha x,$$

where $x(0) = 1$. Here the rate of change is proportional to the state of the system. This is a simple linear equation whose solution is the exponential function $x(t) = e^{\alpha t}$. The only

steady state of this system is $x_{ss} = 0$, and it is unstable.

Although many nonlinear models in biology cannot be solved analytically modellers can rely on numerical methods to obtain accurate approximations to the solutions (Süli and Mayers, 2003). To that end, there are a plethora of analytical methods and tools (eg perturbation methods, bifurcation and stability analysis) that help to understand and approximate the dynamics of a system without having to write an explicit solution (Bender and Orszag, 1999; Kuznetsov, 1998; Seydel, 2010; Strogatz, 1994).

3.2 Reaction and enzyme kinetics

In cell biology many signals are transmitted by means of chemical and biochemical reactions; in this section we introduce a few examples of how these reactions are modelled.

3.2.1 Mass-action kinetics

One of the simplest and most convenient ways of modelling chemical reactions is using *mass-action kinetics* (also known as the law of mass action), first described in 1864 by Waage and Gulberg (1986). Broadly, the law of mass-action states that the rate of reaction of two components in a well-stirred container is proportional to their concentrations. The use of the law of mass action is ubiquitous in biological systems modelling (Chen *et al.*, 2010; Feliu *et al.*, 2010; Gunawardena, 2011; Higham, 2008).
If A and B are two substances that participate in the irreversible reaction

$$A + B \xrightarrow{\alpha} C,$$

the change in their concentration is given by the nonlinear system of equations

$$\begin{aligned}
\frac{\mathrm{d}[A]}{\mathrm{d}t} &= -\alpha[A][B], \\
\frac{\mathrm{d}[B]}{\mathrm{d}t} &= -\alpha[A][B], \\
\frac{\mathrm{d}[C]}{\mathrm{d}t} &= \alpha[A][B],
\end{aligned}$$

where $[A]$, $[B]$ and $[C]$ are the variables of the model and the parameter α is the rate constant of the reaction. The initial conditions of the equations are $[A](0) = A_0$, $[B](0) = B_0$, and $[C](0) = 0$. If the reaction is reversible, eg

$$A + B \underset{\beta}{\overset{\alpha}{\rightleftharpoons}} C,$$

then the equations are

$$\begin{aligned}
\frac{\mathrm{d}[A]}{\mathrm{d}t} &= -\alpha[A][C] + \beta[C], \\
\frac{\mathrm{d}[B]}{\mathrm{d}t} &= -\alpha[A][B] + \beta[C], \\
\frac{\mathrm{d}[C]}{\mathrm{d}t} &= \alpha[A][B] - \beta[C].
\end{aligned}$$

Note that in both sets of equations the negative terms are multiplied by the concentration in whose equations it is found, eg $-\alpha[A][B]$ in the equations for $[A]$ and $[B]$, and $-\beta[C]$ in the equation for $[C]$; this guarantees that the variables in the equations cannot ever become negative, a good reality check for the model.

3.2.2 Enzyme kinetics

Many reactions inside cells are catalysed by enzymes. The typical mode of action of an enzyme is described by the reaction

$$S + E \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} SE \xrightarrow{k_2} P + E, \tag{3.2}$$

where the substrate S reversibly binds to enzyme E to form the substrate-enzyme complex SE, which either dissociates back to S and E or reacts to create a product P. The enzyme is not changed by the reaction and can bind to a new molecule of the substrate to catalyse another reaction.

The variables $[E]$, $[S]$, $[SE]$, and $[P]$ are concentrations of the reactants, the differential

equations that describe the reaction (3.2) (Cornish-Bowden, 2004) are

$$\frac{\mathrm{d}[E]}{\mathrm{d}t} = -k_1[S][E] + k_{-1}[SE] + k_2[SE], \tag{3.3}$$

$$\frac{\mathrm{d}[S]}{\mathrm{d}t} = -k_1[S][E] + k_{-1}[SE], \tag{3.4}$$

$$\frac{\mathrm{d}[SE]}{\mathrm{d}t} = k_1[S][E] - k_{-1}[SE] - k_2[SE], \tag{3.5}$$

$$\frac{\mathrm{d}[P]}{\mathrm{d}t} = k_2[SE], \tag{3.6}$$

with initial conditions:

$$[E](0) = E_0,\ [S](0) = S_0\ [SE](0) = 0,\ [P](0) = 0.$$

Equation (3.6) is decoupled from the rest, ie depends only on $[SE]$ and no other equations depend on it, so we will consider it separately from the other three. Equations (3.3) and (3.5) cancel each other out:

$$\frac{\mathrm{d}[E]}{\mathrm{d}t} + \frac{\mathrm{d}[SE]}{\mathrm{d}t} = 0,$$

which means that $[E] + [SE]$ is always a constant (assuming no gene expression). More specifically, $[E] + [SE] = E_0$ which means that we can eliminate an equation to get

$$\frac{\mathrm{d}[S]}{\mathrm{d}t} = -k_1[S](E_0 - [SE]) + k_{-1}[SE], \tag{3.7}$$

$$\frac{\mathrm{d}[SE]}{\mathrm{d}t} = k_1[S](E_0 - [SE]) - (k_{-1} + k_2)[SE]. \tag{3.8}$$

The units of equations (3.7) and (3.8) are concentration$\times$time^{-1}; for example, if the concentration units are molar and time units are seconds, then the units of the equations are Ms^{-1}.

Nondimensionalisation and the Michaelis-Menten form

It is standard mathematical practice to *nondimensionalise* models, this means dividing the quantities in the model by an appropriate scale (Fowler, 1997; Howison, 2003).

Nondimensional equations and variables are useful because they allow us to know what is small and can be ignored and what has to be kept in a model. We introduce the nondimensional variables $\widehat{[S]}$, $\widehat{[SE]}$, and $\hat{t}$ such that

$$\widehat{[S]} = A[S],\ \widehat{[SE]} = B[SE],\ \hat{t} = Ct. \tag{3.9}$$

The constants A, B, and C are scales with units M (A and B), and s (C) chosen so that the new variables are nondimensional and of order 1 ($O(1)$). Substituting (3.9) in equations (3.7) and (3.8) we obtain

$$\begin{aligned}
\frac{A}{C}\frac{\mathrm{d}\widehat{[S]}}{\mathrm{d}\hat{t}} &= -k_1 A\widehat{[S]}(E_0 - B\widehat{[SE]}) + k_{-1}B\widehat{[SE]},\\
\frac{B}{C}\frac{\mathrm{d}\widehat{[SE]}}{\mathrm{d}\hat{t}} &= k_1 A\widehat{[S]}(E_0 - B\widehat{[SE]}) - B(k_{-1}+k_2)\widehat{[SE]}.
\end{aligned}$$

Choosing scales $A = S_0$, $B = E_0$, and $C = \frac{1}{k_1 E_0}$, and inserting them in the equations and gathering the terms (dropping the hat notation) gives

$$\frac{\mathrm{d}[S]}{\mathrm{d}t} = -(1 - [SE])[S] + (\kappa - \lambda)[SE], \tag{3.10}$$

$$\varepsilon\frac{\mathrm{d}[SE]}{\mathrm{d}t} = [S] - ([S] + \kappa)[SE], \tag{3.11}$$

with $\varepsilon = \frac{E_0}{S_0}$, $\kappa = \frac{k_{-1}+k_2}{S_0 k_1}$, and $\lambda = \frac{k_2}{S_0 k_1}$. The initial conditions now are $[S](0) = 0$ and $[SE](1) = 1$. Usually the initial amount of substrate is much larger than the total amount of enzyme, ie $S_0 \gg E_0$ which means that $\varepsilon \ll 1$. The small parameter in equation (3.11) means that $\varepsilon\frac{\mathrm{d}[SE]}{\mathrm{d}t} \approx 0$, and

$$[SE] \approx \frac{E_0[S]}{\kappa + [S]}. \tag{3.12}$$

When there is much more substrate than enzyme, the formation of the complex $[SE]$ in equation (3.11) reaches equilibrium before there is any substantial decrease in $[S]$, and the expression in equation (3.12) is a good approximation to $[SE]$, that can be used to solve equation (3.10). We can approximate the amount of product of the reaction with:

$$\frac{\mathrm{d}[P]}{\mathrm{d}t} \approx \frac{V_{\max}[S]}{\kappa + [S]}, \tag{3.13}$$

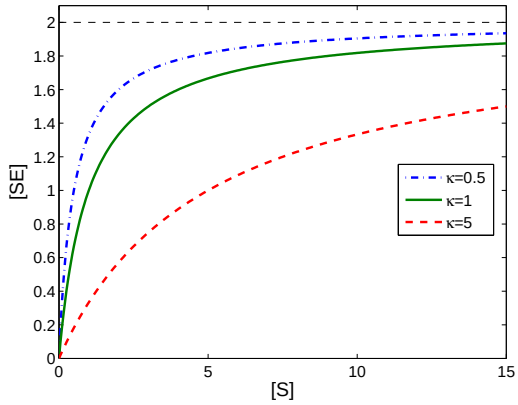

Figure 3.1: Michaelis-Menten reaction rates from equation (3.13) (in arbitrary units) where $V_{\max} = 2$ and $\kappa = 0.5$ (blue dash-dotted line), $\kappa = 1.0$ (green continuous line), and $\kappa = 5.0$ (red dashed line). The horizontal dashed line is the maximum rate of reaction $V_{\max}$.

where $V_{\max} = k_2 E_0$ is the maximum rate of the reaction, and κ is the Michaelis constant. Note that

$$\lim_{[S]\to\infty} \frac{[S]}{\kappa + [S]} = 1,$$

so no matter how much more substrate is added, the reaction rate is limited by the amount of enzyme present in the system. The form in equation (3.12) is called the Michaelis-Menten form and was first proposed by Michaelis and Menten (1913). Figure 3.1 shows an example of the effect of $V_{\max}$ and κ on the reaction rates. Small values of κ indicate that the substrate and the enzyme have high affinity for each other and the Michaelis-Menten term saturates with relatively low quantities of $[S]$ (eg the blue dash-dotted line), larger values indicate lower affinity and demand more substrate to saturate the reaction rate. In this example the maximum rate is given by $V_{max} = 2$, as $[S]$ becomes larger, the rates will get closer to V_{max}.

The Michaelis-Menten approximation of $[SE]$, known as the *quasi-steady-state assumption* (QSSA), is only adequate when $t \gg \varepsilon$. To obtain the solution for earlier times the assumption that $\varepsilon \frac{\mathrm{d}[SE]}{\mathrm{d}t} \approx 0$ no longer holds because $\frac{\mathrm{d}[SE]}{\mathrm{d}t}$ is too large (order $O(\frac{1}{\varepsilon})$), and a different scaling for t must be used (Fowler, 1997; Maini *et al.*, 1991; Segel and Slemrod, 1989). In this work we are mostly interested in the case when $t \gg \varepsilon$, where the Michaelis-Menten form and the QSSA are appropriate.

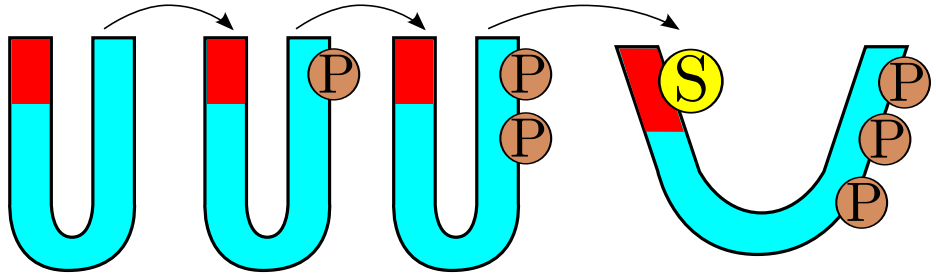

Figure 3.2: Activation of a protein by multiple phosphorylation. When unphosphorylated the protein is in a conformational state that shields its active site (in red) and prevents it from participating in a reaction. After binding three phosphates, a conformational change occurs and the active sites is exposed, thus allowing substrates to bind and other reactions to take place.

3.2.3 The Hill equation

Some enzymes need to bind more than one molecule in order to catalyse a reaction

$$\begin{aligned} S + E &\underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} SE, \\ SE + S &\underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} 2SE, \\ &\vdots \\ (n-1)SE + S &\underset{k_{-n}}{\overset{k_n}{\rightleftharpoons}} nSE \xrightarrow{k_{n+1}} P + E, \end{aligned}$$

and the reaction rates have the shape of a sigmoid rather than the hyperbolic curves shown in Fig. 3.1. For example, Fig. 3.2 shows how a protein undergoes conformational change after binding three phosphate groups, allowing it to expose its active site (in red) and to participate in a reaction. To model such reactions the *Hill equation* is often used. The Hill equation has the form

$$V = \frac{V_{\max}[S]^n}{\kappa^n + [S]^n}, \tag{3.14}$$

where $V_{\max}$ is the maximum reaction rate, κ is the constant that defines the dose of $[S]$ for which $V = \frac{V_{\max}}{2}$ (if $n = 1$ then κ is the Michaelis constant), and n is the *Hill coefficient*. The value of n is an indicator of "cooperativity" between active sites of the enzyme, the higher its value the more cooperative they are, but it does not mean that exactly n molecules

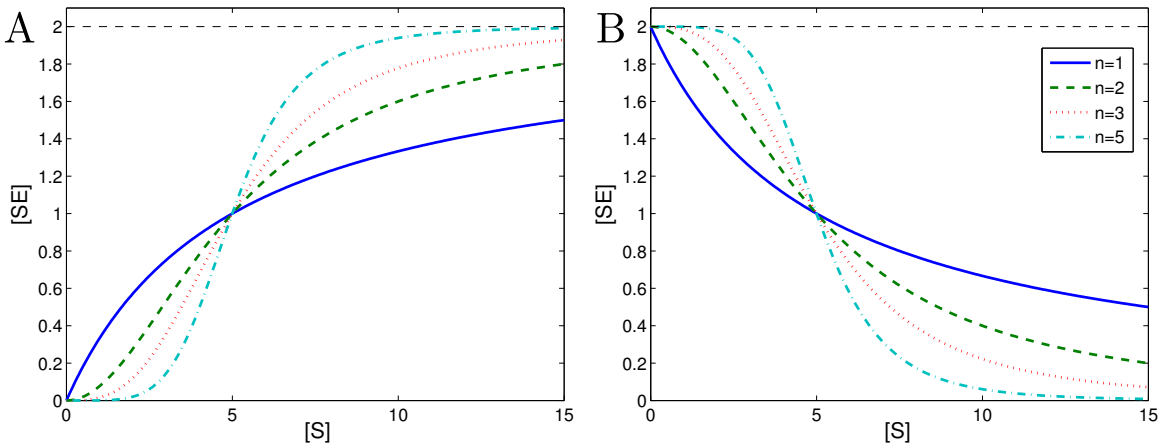

Figure 3.3: **A**: Reaction rates of the Hill expression in equation (3.14) when $V_{\max} = 2$, $\kappa = 5$, and $n = 1$ (dark blue continuous line), $n = 2$ (green dashed line), $n = 3$ (red dotted line), and $n = 5$ (light blue dash-dotted line). **B**: Reaction rates of the inhibitory form in equation (3.15) with the same parameter values as image A.

have to bind the enzyme (Cornish-Bowden, 2004). Often the value of n is a real non-negative number.

Figure 3.3A shows the reaction rates of equation (3.14) with different values of n. When $n = 1$ we observe the Michaelis-Menten curve from Fig. 3.1. For $n > 1$ the response curve is a sigmoid that becomes steeper as n increases.

Hill equations have also been used to model the case when multiple molecules bind to an enzyme to stop it catalysing a reaction:

$$V = \frac{V_{\max}\kappa^n}{\kappa^n + [S]^n}. \tag{3.15}$$

Figure 3.3B shows the reaction rates for different values of n. Again, the larger the n, the more steep the sigmoid becomes.

The two forms of the Hill function from equations (3.14) and (3.15) are used to model not only cooperative enzyme dynamics, but also used to model empirical relationships between components in signalling systems, as we show in the following section.

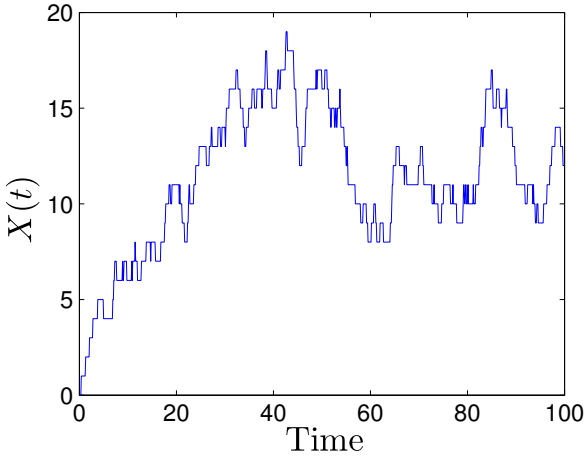

Figure 3.4: Simulation of the production of a molecule given by the stochastic process (3.16) and (3.17), with parameters $\alpha = 1$, $\mu = 0.1$, $dt = 0.1$ and initial condition $X(t) = 0$.

3.2.4 Stochastic models

A different approach to modelling signalling and biochemical processes in the cell is with stochastic models. The cell is an intrinsically noisy environment and no two cells function in exactly the same way, not even cells with identical genotype (Wilkinson, 2009). Instead of dealing with rates of deterministic processes, in stochastic models we consider the probability of occurrence of events (Taylor and Karlin, 1998). Stochastic models can be defined in discrete (eg a discrete-time Markov chain) or continuous time (eg a Wiener process).

Example 3.2.1 (Wilkinson (2009))**.** Let $X(t)$ be the number of copies of a molecule at time t; the number of molecules after a time-step h (ie $X(t+h)$) is determined by the probability of having a new molecule added and the probability of having a molecule degraded, given the value of $X(t)$. In terms of probabilities:

$$P(X(t+h) = X(t) + 1 | X(t)) = \alpha h, \tag{3.16}$$

$$P(X(t+h) = X(t) - 1 | X(t)) = \min(\mu h X(t), 1), \tag{3.17}$$

where α and μ are analogous to rates of production and degradation from ODE models. Figure 3.4 shows the time-course of a single simulation of the process defined by the

probabilities (3.16) and (3.17). The analytical solution of the process is that $X(t)$ is a Poisson-distributed random variable, whose parameter depends on α, μ, and t.

The example we have just given is simple, models with more species interacting in complicated ways require sophisticated theory and intensive computing techniques (Gillespie, 1977; Higham, 2008; Iliopoulos *et al.*, 2010; Milner *et al.*, 1992).

3.3 Cell signalling networks

Cells function by sensing and reacting to their environment; to do so they must interpret external signals and relay them through a chain of components and events that together form a signalling network. Cell signalling networks are formed by the chemical and physical interactions among the different components of the cell that participate in the signalling processes that elicit a response. Models of cellular signalling networks can be constructed with varying levels of detail, usually depending on the data available. A model of a network that includes every single physical and chemical interaction may not be useful at all because the kinetic parameters and species are unlikely to be measurable or identifiable, and investigating its behaviour can be just as complicated as studying the system experimentally. In this section we provide a brief introduction to networks and some examples from cellular network modelling.

3.3.1 Minimal introduction to networks

A network or graph is a mathematical object consisting of a collection of agents (nodes) and their relations (edges) (Gross and Yellen, 2004). An edge may represent chemical or physical interaction, correlation, flux, presence in common places, or communication (Newman, 2003; Stumpf and Wiuf, 2010). The connectivity of a graph (which nodes are connected to which) is known as the topology of the network. If a network has m nodes, the $m \times m$ adjacency matrix of the network $\mathbf{G}$ has nonzero entries for existing edges, and zeros everywhere else. For example, if in a network there is a node between nodes in columns 1 and 4 of $\mathbf{G}$ then $\mathbf{G}(1,4) \neq 0$. In an *undirected network* (Fig. 3.5A) the relationship of the nodes is symmetric, ie the edge from node N_A to node N_B, $\omega_{A,B}$ is equivalent to $\omega_{B,A}$,

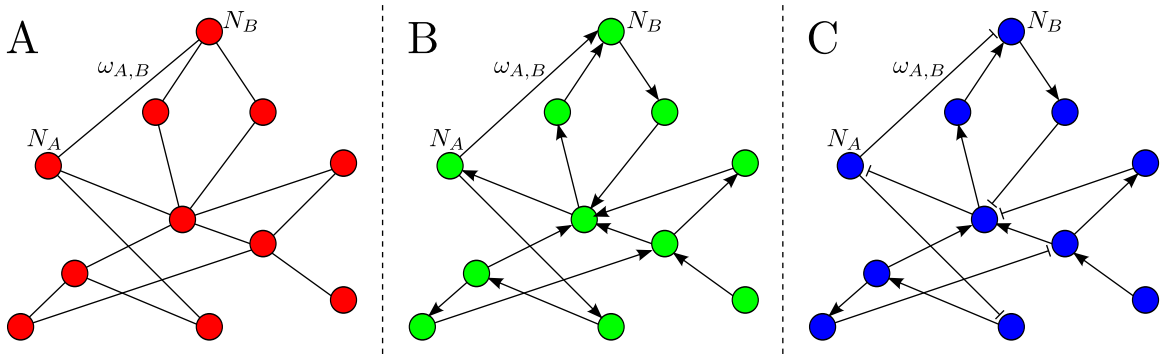

Figure 3.5: Different types of networks. **A:** Undirected network, the relationships between the nodes do not require that edges have a specific origin and end point. **B:** Directed network, the relationship between two nodes is directional, and the edges have a precise origin and destination. **C:** Signed *and* directed network, the relationships among nodes can either be *positive* (edges ending with $\rightarrow$) or *negative* (edges ending with $\dashv$).

the edge from N_B to N_A. The adjacency matrices of undirected networks are symmetric, ie $\mathbf{G} = \mathbf{G}^T$. Examples of undirected relationships can be acquaintanceship of two people in a social network (Traud *et al.*, 2011), strong correlation of two variables in a data set (Onnela *et al.*, 2003), and interaction of proteins in yeast (Schwikowski *et al.*, 2000). A *directed network* (Fig. 3.5B) is a network where the interactions are unidirectional; the existence of an edge from node N_A to node N_B does not imply the reciprocal edge exists, so $\omega_{A,B} \neq \omega_{B,A}$. Adjacency matrices of directed networks are not necessarily symmetric. Examples of directed networks are disease transmission networks (Meyers *et al.*, 2006), and messages on social networks (Huberman *et al.*, 2008). In a *signed network* edges can be described as positive or negative. Figure 3.5C is an example of a signed and directed network where the relationship between nodes N_A and N_B is negative $\omega_{A,B} < 0$; however, not all signed networks are necessarily directed. The entries of the adjacency matrix of a signed network may have positive and negative entries, depending on the sign of the edge. Examples of signed networks include conflict networks (eg amity and enmity) (Wasserman and Faust, 1994), metabolic and gene regulatory networks (Alon, 2007). In addition to direction or sign, the edges on a network can be *weighted* or *unweighted*. The weight of an edge provides additional information about the relationship between two nodes (Newman, 2003). In an unweighted network, all that is known about an edge is whether it exists or not, ie $\omega_{A,B}$ can either be 1 or 0. The entries of $\mathbf{G}$ in an unweighted network are either 1 or

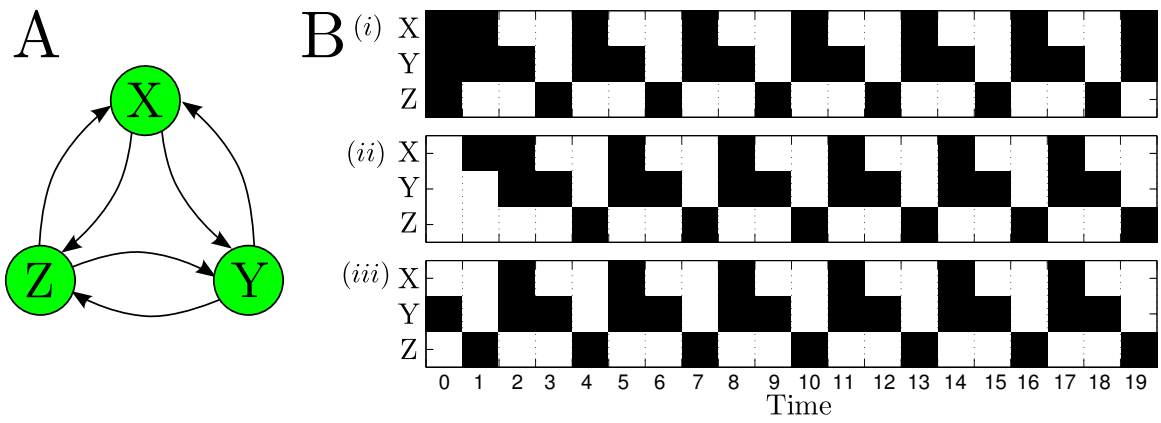

Figure 3.6: **A:** A Boolean network of three variables. **B:** Three simulations of the network where the updates follow rules (3.18)-(3.20), starting from a different states (see text). Each row tracks the state of one variable at 20 discrete time points from 0 to 19; black cells mark a value of 0 and white cells mark a value of 1.

0, whereas in weighted networks they may have other values.

Importantly, in many networks the edges and their characteristics can change in time (Mucha *et al.*, 2010; Palla *et al.*, 2007), so $\omega_{A,B}(t)$ can be a function that represents the time-changing weight of the edge between N_A and N_B.

3.3.2 Boolean network models

In Boolean models the variables can only be in one of two states *on* or *off* usually represented by 1 and 0, respectively (Gershenson, 2004). Boolean networks are formed by boolean variables whose interactions define the rules by which the states of each variable are updated (Kauffman, 1969).

Example 3.3.1. A simple boolean network. In Fig. 3.6A we show a network of three Boolean variables X, Y, and Z, each of which influences and is influenced by the other two (ie the network is directed and fully-connected). In this example by Kauffman (1969)

X_n, Y_n, and Z_n denote the state of the variables at time n, the rules for updating are:

$$X_{n+1} = Z_n \wedge (\neg Y_n), \tag{3.18}$$

$$Y_{n+1} = X_n \wedge Z_n, \tag{3.19}$$

$$Z_{n+1} = \neg(X_n \wedge (\neg Y_n)), \tag{3.20}$$

where $\wedge$, and $\neg$ are the standard logic operators *and* and *not*, respectively. On Fig. 3.6B we show three simulations of the boolean network in Fig. 3.6A, using rules (3.18)-(3.20) and starting from three different initial conditions:

$$(i) \quad [X_0\ Y_0\ Z_0] = [0\ 0\ 0],$$
$$(ii) \quad [X_0\ Y_0\ Z_0] = [1\ 1\ 1],$$
$$(iii) \quad [X_0\ Y_0\ Z_0] = [1\ 0\ 1].$$

On all simulations, the network's state settled into oscillations with period-three, and simulation (i) is one time-step ahead of (ii) and (iii), who are synchronised from the second iteration. The recurring states of the system are:

$$[0\ 0\ 1],\ [1\ 0\ 1],\ \text{and}\ [1\ 1\ 0].$$

More sophisticated boolean models and networks may have nontrivial dynamics, include randomness, complex topologies, and asynchronous (non-simultaneous) updating of the nodes (Aldana, 2003; Gershenson, 2003, 2004; Kauffman, 1969; Saadatpour *et al.*, 2010).

3.3.3 ODE network models

Differential equation models of cell signalling networks typically have one equation per node which describes the change in the state of the node, or a quantity that describes it. The edges are the interactions that form the equations. (Sauro, 2009). Below we show two examples of ODE network models that display different behaviour.

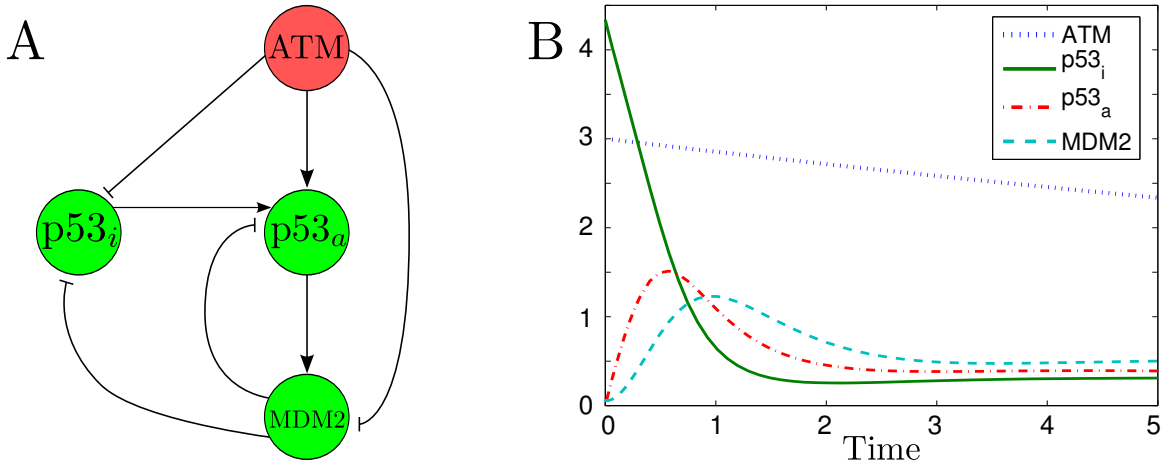

Figure 3.7: **A:** Regulatory network of p53 upon DNA damage (Brewer, 2006). The network is weighted, directed, signed, and time-dependent. **B:** Time course solutions of the model in equations (3.21)-(3.24) using the parameters stipulated by Brewer *et al.* (2008): $\beta_{10} = 0.05$, $\alpha_{20} = 0.52$, $k_1 = 1.42$, $\beta_{20} = 0.041$, $k_2 = 0.39$, $\alpha_{40} = 0.01$, $k_3 = 2.5$, $k_4 = 0.75$, and $\beta_{40} = 0.18$, and initial conditions $a(0) = 3$, $z(0) = 4.33$, $x(0) = 0$ $y(0) = 0.05$ (the unstimulated steady state for all variables except $a(0)$ which has been increased to simulate the DNA damage stimulus).

Example 3.3.2. Regulation of p53 expression. The protein p53 is crucial in preventing the proliferation of cells with mutated or damaged DNA by inducing cell cycle arrest and DNA repair, or inducing cell suicide. Mutations or alterations of normal p53 function often lead to uncontrolled cell proliferation and cancer (Weinberg, 2007). In Figure 3.7A we show the p53 regulatory network. DNA damage induces the activation of the protein ataxia telangiectasia mutated (ATM) which phosphorylates inactive p53. Active p53 induces the expression of the murine double minute (MDM2) oncogene, which accelerates the degradation of both active and inactive p53. In addition to activating p53, ATM also helps the degradation of active MDM2. The p53 regulatory network we have described has been

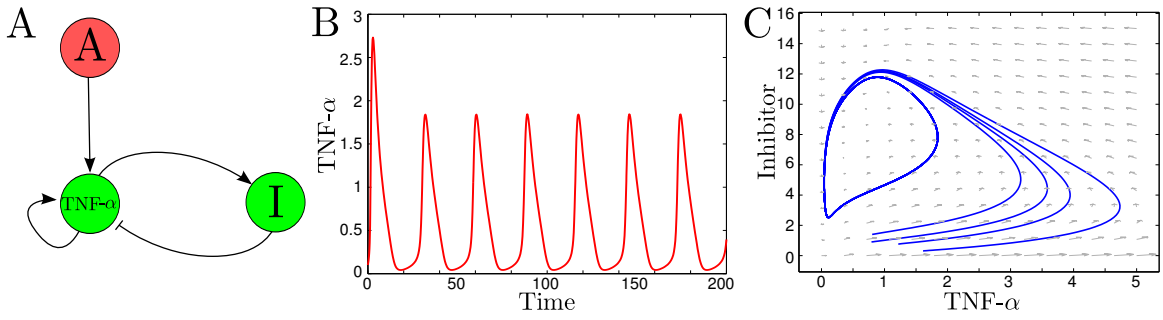

Figure 3.8: **A:** Network representation of the control of TNF-α expression in rabbits following corneal transplant. The network is formed by nodes A (the external signal), $TNF-\alpha$ (the concentration of the tumour necrosis factor), and I (the inhibitor pool); it is also signed, weighted and time-dependent. **B:** Sustained oscillations of TNF-α given by equations (3.25) and (3.26) using parameters $A = 21$, $B = 0$, $C = 5$, $D = 0.2$, $E_1 = 0.1$, $E_2 = 0.001$, $F = 1$, and $n = 2$, and initial conditions $u(0) = 0.1$ and $v(0) = 0.1$. **C:** Trajectories of the solutions of equations (3.25) and (3.26) in the phase plane from different initial conditions.

modelled with ODEs using mass-action kinetics by Brewer (2006), Brewer *et al.* (2008):

$$\frac{\mathrm{d}a}{\mathrm{d}t} = -\beta_{10}a, \tag{3.21}$$

$$\frac{\mathrm{d}z}{\mathrm{d}t} = \alpha_{20} - (k_1 y + k_2 a + \beta_{20})z, \tag{3.22}$$

$$\frac{\mathrm{d}x}{\mathrm{d}t} = k_2 a z - (k_1 y + \beta_{20})x, \tag{3.23}$$

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \alpha_{40} + k_3 x - (k_4 a + \beta_{40})y, \tag{3.24}$$

where a is active ATM, z; inactive p53, x; active p53, and y; active MDM2. Figure 3.7B shows the time course behaviour of the model in equations (3.21)-(3.24).

Example 3.3.3. Cytokine oscillations. Tumour necrosis factor-α (TNF-α) is a cytokine involved in inflammatory response. In rabbits that have rejected corneal transplant TNF-α has been reported to oscillate (Rayner *et al.*, 2000). The regulation of TNF-α is by no means trivial given that a number of other cytokines can modulate its expression (eg IL-10, TGB-β), and TNF-α itself can amplify its own effects. A simplified network depiction of the TNF-α regulatory dynamics is shown in Fig. 3.8A. An external signal (node A) promotes the expression of TNF-α which has a self-loop (promotes its own expression) and activates

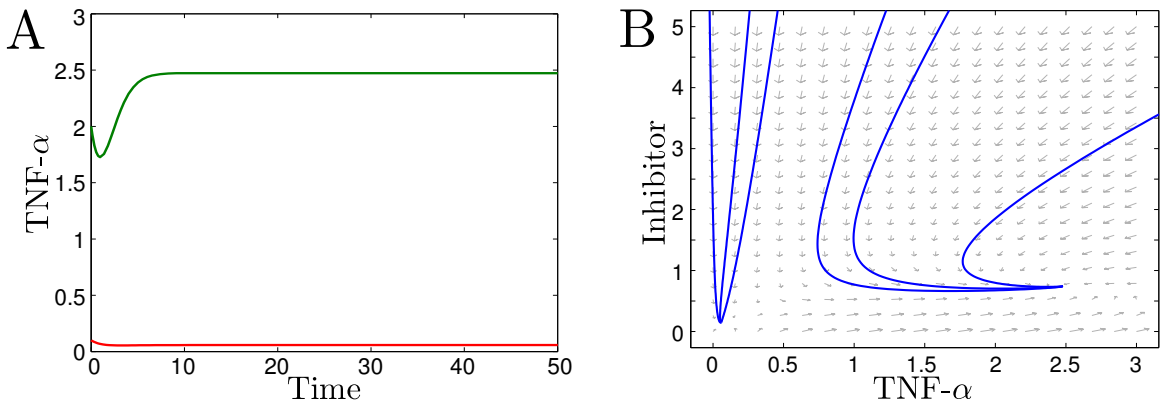

Figure 3.9: **A:** Temporal dynamics of TNF-α. Using parameters $A = 5$, $B = 0$, $C = 1$, $D = 1$, $E_1 = 0.1$, $E_2 = 0.1$, $F = 1$, and $n = 2$, the steady state depends only on the initial conditions. Using $u(0) = 0.1$ and $v(0) = 0.1$ gives a low level of TNF-α (red line), using $u(0) = 2$ and $v(0) = 2$; a high level of TNF-α (green line). **B:** Phase plane with two sinks (stable steady states), the system will go to one or the other depending on where it begins.

its inhibitor pool (node I). This TNF-α regulatory network has been modelled by Chan *et al.* (1999) using the ODEs:

$$\frac{\mathrm{d}u}{\mathrm{d}t} = A\frac{u^n + E_1^n}{u^n + 1}\frac{1}{1+v} - u, \tag{3.25}$$

$$\frac{\mathrm{d}v}{\mathrm{d}t} = B + C\frac{u + E_2}{u + F} - Dv, \tag{3.26}$$

where $u(t)$ denotes the concentration of TNF-α at time t, $v(t)$ is the concentration of the inhibitor pool, and A is the external stimulus; both equations are nondimensional. The change in u is driven by the external stimulus A, TNF-α's auto-activation, inhibition by v, and linear inactivation. The change in v is driven by constant production B, activation by u and linear inactivation. In this case, the external stimulus is constant so $\frac{\mathrm{d}A}{\mathrm{d}t} = 0$. Equations (3.25) and (3.26) and their dynamics have a diverse range of behaviour including mono and multistability, excitability, and oscillations (Chan *et al.*, 1999). For example in Fig. 3.8B we show sustained oscillations of TNF-α, which resemble experimental observations (Rayner *et al.*, 2000). Figure 3.8C shows how different trajectories on the TNF-α $\times$ inhibitor phase plane settle into a steady cycle. Changing the values of the parameters with which we produced Figs. 3.8B and 3.8C results in a radically different observed behaviour (ie a bifurcation). In Figs. 3.9A and 3.9B we solved the same two

equations (3.25) and (3.26) with a different set of parameters which gives two stable steady states instead of the only one observed previously (Chan *et al.*, 1999).

The model of TNF-α regulation in equations (3.25) and (3.26) shows how even simplifications of signalling processes of cells can give rise to rich dynamic landscapes that can be used to guide further experimentation.

3.4 Discussion

In this chapter the basic modelling tools used in mathematical modelling of cell signalling systems have been introduced. Brief descriptions and examples of stochastic and boolean modelling were given, with special emphasis on differential equation models. These concepts are used extensively in the next chapters to analyse activation cascades and to construct a model of stomatal closure.

Chapter 4

Evolutionary Monte Carlo optimisation with local search acceleration

4.1 Introduction

The increasing drive towards quantitative technologies in Biology has brought with it a renewed interest in the modelling of biological systems. As shown in Chapters 3, 5, and 7, models of biological systems and other complex phenomena are generally nonlinear with uncertain parameters, many of which are often unknown and/or unmeasurable (Alon, 2007; Edelstein-Keshet, 1988). Crucially, the values of the parameters dictate not only the quantitative but also the qualitative behaviour of such models (Brown and Sethna, 2003; Strogatz, 1994). A fundamental task in quantitative and systems biology is to use experimental data to infer parameter values that minimise the discrepancy between the behaviour of the model and experimental observations. The parameters thus obtained can then be cross-validated against unused data before employing the fitted model as a predictive tool (Alon, 2007). Ideally, this process could help close the modelling-experiment loop by: suggesting specific experimental measurements; identifying relevant parameters to be measured; or discriminating between alternative models (Gutenkunst *et al.*, 2007; Toni and Stumpf, 2010; Yates *et al.*, 2001).

The problem of parameter estimation and data fitting is classically posed as the minimisation of a cost function (ie the error) (Gershenfeld, 1999). In the case of

overdetermined linear systems with quadratic error functions, this problem leads to least-square solutions, convex optimisations that can be solved efficiently and globally based on the singular value decomposition of the covariance matrix of the data (Lawson and Hanson, 1995). However, data fitting in nonlinear systems with small amounts of data remains difficult, as it usually leads to non-convex optimisations with several local minima (Brewer *et al.*, 2008).

A classic case in biological modelling is the description of the time evolution of a system through ordinary differential equations (ODEs), usually based on mechanistic functional forms such as the ones introduced as examples in Chapter 3, the models of activation cascades from Chapter 5, and the models of stomatal closure developed in Chapter 7. Typically, optimal parameters of the nonlinear ODEs must be inferred from experimental time courses but the associated optimisation is far from straightforward. Standard optimisation techniques that require an explicit cost function are unsuitable for this problem due to the difficulty to obtain full analytical solutions for nonlinear ODEs (Brown and Sethna, 2003; Chen *et al.*, 2010; Papachristodoulou and Recht, 2007). Spline-based methods, which approximate the solution though an implicit integration of the differential equation (Brewer *et al.*, 2008), require linearity in the parameters and are therefore not applicable to models with nonlinear parameter dependencies, eg Michaelis-Menten and Hill kinetics.

Implicit techniques, such as direct search methods (Powell, 1998), Simulated Annealing (Kirkpatrick *et al.*, 1983), Evolutionary Algorithms (Mitchell, 1997; Runarsson and Yao, 2000) or Sequential Monte Carlo (Sisson *et al.*, 2007), do not require an explicit cost function. However, if (as is usually the case) the cost function is a complicated (hyper)surface in parameter space with many local minima. Gradient and direct search methods tend to get trapped in local minima due to their use of local information. Although still a local method, Simulated Annealing alleviates some of the problems related to local minima through the use of stochasticity. However, its success comes at the cost of a high computational overhead and slow convergence, without guarantees of finding the global minimum.

Instead of an optimisation based on local criteria, Evolutionary Algorithms (EA) produce an ensemble of possible answers and evolve them globally through random

mutation and cross-over followed by ranking and culling of the worst solutions (Mitchell, 1997; Runarsson and Yao, 2000; Schwefel, 1995). This heuristic has been shown to provide an efficient protocol for parameter fitting in the life sciences (Moles *et al.*, 2003; Zi and Klipp, 2006). However, EA methods can be inefficient when the feasible region in parameter space is too large, a case typical of models with large uncertainty in the parameters.

Probabilistic methods, such as Sequential Monte-Carlo (SMC) (Sisson *et al.*, 2007), propose a different conceptual framework. Rather than finding a *unique* optimal parameter set, SMC maps a prior probability distribution of the parameters onto a posterior constructed from samples with low errors until reaching a converged posterior. Recently, SMC has been combined with Approximate Bayesian Computation (ABC) and applied to data fitting and model selection (Toni *et al.*, 2009). However, methods such as ABC-SMC are not only computationally expensive but also require the starting prior to include the *true* value of the parameters. This requirement dents its applicability to many biological models, in which not even the order of magnitude of the parameters is known. In that case, the support of the starting priors must be made overly large (leading to extremely slow convergence) in order to avoid the risk of excluding the true parameter value from the search space.

In this chapter, we present a novel optimisation algorithm for data fitting that takes inspiration from EA, SMC and direct search optimisation. This method iterates and refines samples from a probability distribution of the parameters in a 'squeeze-and-breathe' sequence. At each iteration the probability distribution is 'squeezed' by the consecutive application of local optimisation followed by ranking and culling of the local optima. The parameter distribution is then allowed to 'breathe' through a random update from a historical prior that includes the union of all past supports of the solutions (Fig. 4.1). This iteration proceeds until convergence of the distribution of solutions and their average error. A key, distinctive feature of this algorithm is the accelerated step-to-step convergence through a combination of local optimisation and of culling of local solutions. Importantly, the method can also find parameters that lie outside of the range of the initial prior, and can deal with parameter values across several orders of magnitude.

4.2 Algorithm

We now provide definitions and a full description of this algorithm and give examples of its applicability to different biological models of interest.

4.2.1 Formulation of the problem

Let $\mathbf{X}(t) = [x_1(t), \ldots, x_d(t)]$ denote the state of a system with d variables at time t. The time evolution of the state is described by a system of (possibly nonlinear) ODEs:

$$\dot{\mathbf{X}} = f(\mathbf{X}, t; \boldsymbol{\theta}). \tag{4.1}$$

Here $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_N]$ is the vector of N parameters of our model.

The experimental data set is formed by M observations of some of the variables of the system:

$$\mathcal{D} = \left\{ \widetilde{\mathbf{X}}(t_i) \,|\, i = 1, \ldots, M \right\}, \tag{4.2}$$

where $\widetilde{\mathbf{X}}(t_i)$ corresponds to the real value of the system plus observational error. Ideally, $M > 2N + 1$ since $2N + 1$ experiments are enough for unequivocal identification of an ODE model with N parameters when no measurement error is present (Sontag, 2002).

The *cost function* to be minimised is the error of the model given the data:

$$E_{\mathcal{D}}(\boldsymbol{\theta}) = \sum_{i=1}^{M} \left|\left| \mathbf{X}(t_i; \boldsymbol{\theta}) - \widetilde{\mathbf{X}}(t_i) \right|\right|, \tag{4.3}$$

where $||\cdot||$ is a relevant vector norm. A standard choice is the Euclidean norm (or 2-norm) which corresponds to the sum of squared errors:

$$E^{(2)}_{\mathcal{D}}(\boldsymbol{\theta}) = \sum_{i=1}^{M} \sum_{j=1}^{d'} \left(X_j(t_i; \boldsymbol{\theta}) - \widetilde{X}_j(t_i) \right)^2, \tag{4.4}$$

where we assume that d' variables are observed. The cost function $E_{\mathcal{D}} : \mathbb{R}^N \to \mathbb{R}_+$ maps a N-dimensional parameter vector onto its corresponding error, thus quantifying how far the data and the model predictions are for that particular parameter set.

The aim of the data fitting procedure is to find the parameter vector $\boldsymbol{\theta}^{**}$ that minimises the error globally subject to restrictions dictated by the problem of interest:

$$\boldsymbol{\theta}^{**} = \min_{\boldsymbol{\theta}} E_{\mathcal{D}}(\boldsymbol{\theta}), \quad \text{subject to constraints on } \boldsymbol{\theta}. \tag{4.5}$$

4.2.2 Definitions

- *Data set:* $\mathcal{D}$, a set of M observations, as defined in Eq. (4.2).
- *Parameter set:* $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_N] \in \mathbb{R}_+^N$. Due to the nature of the models considered, $\theta_i \geq 0, \ \forall i$.
- *Objective function:* $E_{\mathcal{D}}(\boldsymbol{\theta})$, the error function to be minimised given the data set $\mathcal{D}$, as defined in Eq. (4.4).
- *Set of local minima of* $E_{\mathcal{D}}(\boldsymbol{\theta})$*:*

  $$\mathbb{M} = \{\boldsymbol{\theta}^* \mid E_{\mathcal{D}}(\boldsymbol{\theta}^*) \ \leq \ E_{\mathcal{D}}(\boldsymbol{\theta}), \forall \boldsymbol{\theta} \in \mathcal{N}(\boldsymbol{\theta}^*)\},$$

  where $\mathcal{N}(\boldsymbol{\theta}^*)$ is a neighbourhood of $\boldsymbol{\theta}^*$.
- *Global minimum of* $E_{\mathcal{D}}(\boldsymbol{\theta})$*:* $\boldsymbol{\theta}^{**}$, a parameter set such that $E_{\mathcal{D}}(\boldsymbol{\theta}^{**}) \leq E_{\mathcal{D}}(\boldsymbol{\theta}), \forall \boldsymbol{\theta}$ in the feasible region. Clearly, $\boldsymbol{\theta}^{**} \in \mathbb{M}$.
- *Local minimisation mapping:* $L : \mathbb{R}_+^N \to \mathbb{M}$. Local minimisation maps $\boldsymbol{\theta}$ onto a local minimum: $L(\boldsymbol{\theta}) = \boldsymbol{\theta}^* \in \mathbb{M}$.
- *Ranking and culling of local minima*: $\{\boldsymbol{\theta}^\dagger\}_1^B = \mathcal{RC}_B\left(\{\boldsymbol{\theta}\}_1^J\right)$. This operation ranks J parameter sets and selects the B parameter sets with the lowest value of $E_{\mathcal{D}}$.
- *Joint probability distributions of the parameters at iteration* k*:* $\pi_k(\boldsymbol{\theta})$ (prior) and $\varpi_k(\boldsymbol{\theta})$ (posterior).
- *Marginal probability distribution of the* i^{th} *component of* $\boldsymbol{\theta}$: For instance,

  $$\pi(\theta_i) = \int \pi(\boldsymbol{\theta}) \prod_{r \neq i} \mathrm{d}\theta_r.$$

- *Historical prior at iteration k:* $\zeta_k(\boldsymbol{\theta}) = \prod_{i=1}^{N} \zeta_k(\theta_i)$ where

$$\zeta_k(\theta_i) \sim U\left(\min\left(\mathfrak{Z}_k(\theta_i)\right), \max\left(\mathfrak{Z}_k(\theta_i)\right)\right). \tag{4.6}$$

  Here $U(a, b)$ is a uniform distribution with support in $[a, b]$ and $\mathfrak{Z}_k(\theta_i) = \zeta_{k-1}^{-1} \cup \varpi_k^{-1}$ is the union of the supports of $\varpi_k(\theta_i)$ and $\zeta_{k-1}(\theta_i)$.

- *Update of the prior at iteration k:* $\pi_k(\boldsymbol{\theta}) = \prod_{i=1}^{N} \pi_k(\theta_i)$ with

$$\pi_k(\theta_i) \sim p_m \varpi_k(\theta_i) + (1 - p_m)\zeta_k(\theta_i), \tag{4.7}$$

  that is, a convex mixture of the posterior and the historical prior with weight p_m, from which a new population is sampled in iteration $k + 1$.

- *Re-population:* Obtain a new population of J random points simulated from the prior $\pi_{k-1}(\boldsymbol{\theta})$.

- *Convergence criterion for the error:* The difference between the means of the errors of the posteriors in consecutive iterations is smaller than the pre-determined tolerance:

$$\phi_k = \overline{E_{\mathcal{D}}(\varpi_{k-1}(\boldsymbol{\theta}))} - \overline{E_{\mathcal{D}}(\varpi_k(\boldsymbol{\theta}))} < Tol. \tag{4.8}$$

- *Convergence criterion for the empirical distributions:* The samples of the posteriors in consecutive iterations are indistinguishable at the 5% significance level according to the nonparametric Mann-Whitney rank sum test (Mood *et al.*, 1974):

$$\mathcal{MW}\left(\varpi_k(\boldsymbol{\theta}), \varpi_{k-1}(\boldsymbol{\theta})\right) = 0. \tag{4.9}$$

4.2.3 Description of the algorithm

Algorithm 4.1 presents the pseudo-code for the method using the definitions above. The iterations produce progressively more refined distributions of the parameter vector. At each iteration k, a population simulated from the prior distribution $\pi_{k-1}(\boldsymbol{\theta})$ is locally minimised followed by ranking and culling of the local minima to create a posterior distribution $\varpi_k(\boldsymbol{\theta})$

Algorithm 4.1 Squeeze-and-Breathe optimisation.

Set running parameters of algorithm:
$B, J \in \mathbb{N}$, $p_m \in [0,1]$, ***Tol***
Choose initial priors $\pi_0(\boldsymbol{\theta})$ and $\zeta_0(\boldsymbol{\theta})$.
Set $\mathcal{H}_0 = \emptyset$ and $k \leftarrow 1$.
repeat
  Let $\mathcal{H}_k = \mathcal{H}_{k-1}$.
  Simulate J points from $\pi_{k-1}(\boldsymbol{\theta})$ through re-population.
  for $\ell = 1 \rightarrow J$ **do**
    Obtain local minimum $\boldsymbol{\theta}^*_\ell = L(\boldsymbol{\theta}_\ell)$.
    Store the pair $[\boldsymbol{\theta}^*_\ell, E_{\mathcal{D}}(\boldsymbol{\theta}^*_\ell)]$ in $\mathcal{H}_k$.
  end for
  Rank and cull the set of local minima:
  $\mathcal{H}_k = \mathcal{RC}_B\left(\mathcal{H}_k\right)$
  Define the posterior $\varpi_k(\boldsymbol{\theta})$ from the sample $\mathcal{H}_k$.
  Update $\zeta_k(\boldsymbol{\theta})$ from $\zeta_{k-1}(\boldsymbol{\theta})$ and $\varpi_k(\boldsymbol{\theta})$.
  Update the prior:
  $\pi_k(\boldsymbol{\theta}) \sim p_m \varpi_k(\boldsymbol{\theta}) + (1-p_m)\zeta_k(\boldsymbol{\theta})$.
  $k \leftarrow k+1$.
until $\phi_k < Tol$ and $\mathcal{MW}\left(\varpi_k(\boldsymbol{\theta}), \varpi_{k-1}(\boldsymbol{\theta})\right) = 0$

(squeeze step). This distribution is then combined with an encompassing historical prior to generate the updated prior $\pi_k(\boldsymbol{\theta})$ (breathe step). The iteration loop terminates when the difference between the mean errors of consecutive posteriors (ϕ_k) is smaller than the tolerance and the samples of the posteriors are indistinguishable. We now explain these steps in detail aided by Fig. 4.1 and the BPM model (see Sec. 4.3.1).

1. *Formulation of the optimisation:* The data set $\mathcal{D}$ and the model equations parametrised by $\boldsymbol{\theta}$ allow us to define an error function $E_{\mathcal{D}}(\boldsymbol{\theta})$ whose global minimum corresponds to the best model.

   In our illustrative example, the BPM model (4.10) has the parameter vector $\boldsymbol{\theta} = [\alpha, \beta]$ and the error function is depicted in Fig. 4.1A. Typically, the global optimisation on a rugged landscape of a function like this is computationally hard.

2. *Initialisation:*

   - Set the running parameters of the algorithm: the size of the simulated population, J; the size of the surviving population after culling, B; the update

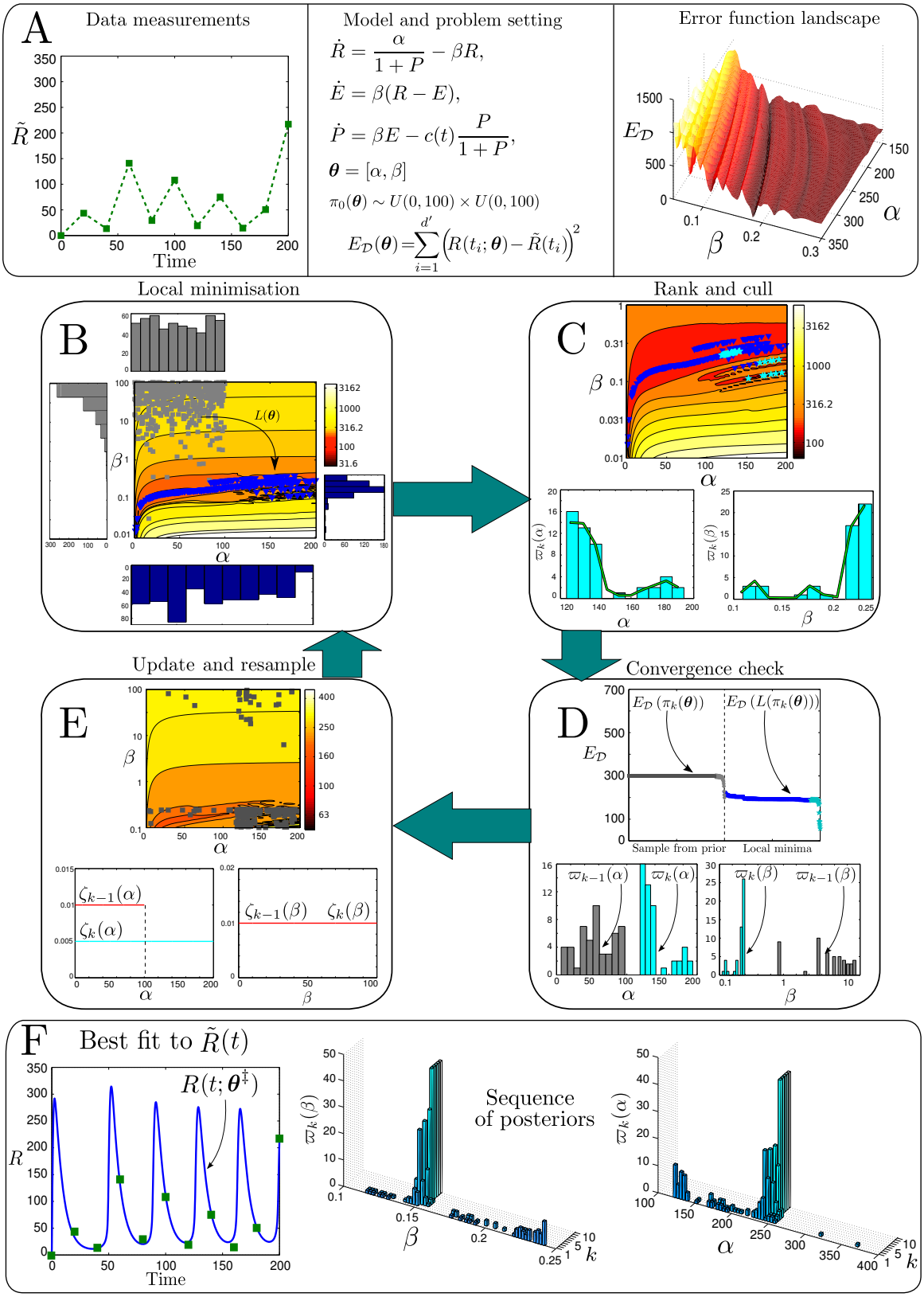
A
Data measurements
Time
Model and problem setting
$\dot{R} = \frac{\alpha}{1+P} - \beta R,$
$\dot{E} = \beta(R - E),$
$\dot{P} = \beta E - c(t)\frac{P}{1+P},$
$\boldsymbol{\theta} = [\alpha, \beta]$
$\pi_0(\boldsymbol{\theta}) \sim U(0,100) \times U(0,100)$
$E_{\mathcal{D}}(\boldsymbol{\theta}) = \sum_{i=1}^{d'} \left(R(t_i;\boldsymbol{\theta}) - \tilde{R}(t_i)\right)^2$
Error function landscape
Local minimisation
B
$L(\boldsymbol{\theta})$
Rank and cull
C
Update and resample
E
Convergence check
D
$E_{\mathcal{D}}(\pi_k(\boldsymbol{\theta}))$
$E_{\mathcal{D}}(L(\pi_k(\boldsymbol{\theta})))$
Sample from prior
Local minima
F
Best fit to $\tilde{R}(t)$
$R(t;\boldsymbol{\theta}^{\ddagger})$
Time
Sequence of posteriors

Figure 4.1 *(preceding page)*: Steps of Algorithm 4.1 exemplified through the BPM model (4.10). **A**: The problem is defined by the data set, the model and the error function to be minimised. Note the rugged landscape of the error function in the parameter plane (α, β), with many local minima. **B**: In the first iteration, we simulate J points in parameter space from the uniform initial prior $\pi_0(\boldsymbol{\theta})$ (squares in plot, top and left histograms) which are then minimised locally with a Nelder-Mead algorithm $L(\boldsymbol{\theta})$ (triangles in plot, bottom and right histograms). The local optimisation aligns the parameter sets onto the level curves of $E_{\mathcal{D}}$. **C**: The B best local minima (top, light stars) are selected and considered to be samples from the posterior distribution (bottom, light histograms). **D**: Convergence of the error of the samples (top plot on the right, B lowest minima are the light stars) and of the posterior distributions (bottom, lighter histograms) are checked against the errors of the sample (top plot on the left) and the priors (bottom, darker histograms). **E**: If convergence is not achieved, the historical prior is updated (previous historical prior in bold, updated in light) and a new set of J points are simulated from the posterior with probability p_m and from the historical prior with probability $1 - p_m$ (squares in plot). This new sample is fed back to the local minimisation step **(b)**. **F**: The algorithm stops when convergence is reached (after nine iterations, in this case) providing an optimal parameter set $\boldsymbol{\theta}^{\ddagger}$ and a time course (top) and the sequence of optimised posteriors at each iteration (bottom).

probability, p_m; and the tolerance, Tol.

In this example, $J = 500$, $B = 50$, $p_m = 0.95$ and $Tol = 10^{-5}$.

- Choose $\pi_0(\boldsymbol{\theta})$, the initial prior distribution of the parameter vector.

  In this case, we take α and β to be independent and uniformly distributed: $\pi_0(\boldsymbol{\theta}) \sim U(0, 100) \times U(0, 100)$.

- Initialise $\zeta_0(\boldsymbol{\theta}) = \pi_0(\boldsymbol{\theta})$, the historical prior of the parameters.

- Simulate J points from $\pi_0(\boldsymbol{\theta})$ to generate the initial sample $\{\widehat{\boldsymbol{\theta}}_0\}_1^J$.

3. *Iteration (step k):* Repeated until termination criterion is satisfied. Figure 4.1B-E shows the first iteration of our method applied to the BPM example.

   (a) *Local minimisation:* Apply local minimisation to the simulated parameters from the 'prior' $\{\widehat{\boldsymbol{\theta}}_{k-1}\}_1^J$ and map them onto local minima of $E_{\mathcal{D}}(\boldsymbol{\theta})$ to generate $\{L(\widehat{\boldsymbol{\theta}}_{k-1})\}_1^J \in \mathbb{M}$.

   Here we use the Nelder-Mead simplex method (Nelder and Mead, 1965), though others can be used. Figure 4.1B shows the simulated points from $\pi_0(\boldsymbol{\theta})$ (squares in plot) and its corresponding histograms (top and left). After local

minimisation, this sample is mapped onto the dark triangles in Fig. 4.1B (dark histograms bottom and right). Note how the local minima align with the level curves of $E_{\mathcal{D}}$ with a markedly different distribution to the uniform prior. Note also that many of the optimised values of α lie outside the range of the prior $(0, 100)$ and are now distributed over the interval $(0, 200)$. On the other hand, the values of β have collapsed inside $(0, 1)$.

(b) *Ranking and culling:* Rank the $J + B$ local minima from the $k - 1$ and k iterations, select the B points with the lowest $E_{\mathcal{D}}$ and cull (discard) the rest:

$$\mathcal{RC}_B \left(\{L(\widehat{\boldsymbol{\theta}}_{k-1})\}_1^J \cup \{\widehat{\boldsymbol{\theta}}_{k-1}^{\dagger}\}_1^B \right) = \{\widehat{\boldsymbol{\theta}}_k^{\dagger}\}_1^B.$$

We consider $\{\widehat{\boldsymbol{\theta}}_k^{\dagger}\}_1^B$ to be a sample from the optimised ('posterior') distribution, $\varpi_k(\boldsymbol{\theta})$ and we denote the best parameter vector of this set as

$$\boldsymbol{\theta}_k^{\ddagger} = \min_{E_{\mathcal{D}}} \left(\{\widehat{\boldsymbol{\theta}}_k^{\dagger}\}_1^B \right).$$

The $B = 50$ best parameter sets are shown (light stars in plot) in Fig. 4.1C (bottom histograms).

(c) *Termination criterion:* Check that the difference between the mean errors of the consecutive optimised samples is smaller than the tolerance: $\phi_k \leq Tol$. We also gauge the 'convergence' of the posteriors through the Mann-Whitney (MW) test to determine if the samples from consecutive posteriors are distinguishable:

$$\mathcal{MW}(\varpi_{k-1}(\boldsymbol{\theta}), \varpi_k(\boldsymbol{\theta})) \equiv \mathcal{MW} \left(\{\widehat{\boldsymbol{\theta}}_{k-1}^{\dagger}\}_1^B, \{\widehat{\boldsymbol{\theta}}_k^{\dagger}\}_1^B \right),$$

where $\mathcal{MW}$ is a 0-1 flag. The MW test gives additional information about the change of the optimised posteriors from one iteration to the next.

Figure 4.1D shows the convergence check for the first iteration of the BPM model: (i) top, errors of the sampled prior (left) with errors of the local minima (right) and the B surviving points (light stars); (ii) bottom, histograms of the prior and the posterior. Clearly, in this iteration neither the error nor the

distributions have converged so the algorithm does not stop, and a new iteration must be performed.

(d) *Update of historical prior and generation of new sample:* If convergence is not achieved, update the historical prior $\zeta_k(\boldsymbol{\theta})$ as a uniform distribution over the union of the supports of the existing historical prior and the calculated posterior (4.6). Equivalently, the support of the historical prior extends over the union of the sequence of all historical priors $\{\zeta_0(\boldsymbol{\theta}), \ldots, \zeta_{k-1}(\boldsymbol{\theta})\}$ and of all posteriors $\{\varpi_1(\boldsymbol{\theta}), \ldots, \varpi_k(\boldsymbol{\theta})\}$.

As shown in Fig. 4.1(e) for the BPM example, the marginal of the historical prior for α is expanded to $U(0, 200)$, since the optimised parameter sets have reached values as high as 200. Meanwhile, the β marginal of the historical prior remains unchanged as $U(0, 100)$ because there has been no expansion of the support.

The historical prior is used to mutate the updated prior before the next iteration by constructing a weighted mixture of the posterior and the historical prior with weight p_m, as shown in (4.7). We re-populate from this updated prior by simulating from the posterior with probability $p_m = 0.95$ and from the historical prior with probability $(1 - p_m)$ to generate the new sample $\{\widehat{\boldsymbol{\theta}}_k\}_1^J$ and iterate back.

Figure 4.1(e) shows the sample of J points simulated from the new prior. The α-components of most points are between 100 and 200 and the β-components are between 0.1 and 1.0, but there are a few that lie outside the support of the posterior. The process in panels (b), (c), (d), and (e) of Fig. 4.1 is iterated for this new population of points.

4. *Output of the algorithm:* When the convergence criteria have been met, the iteration stops at iteration k^* and the minimum of this last iteration, $\boldsymbol{\theta}_{k^*}^{\ddagger}$, is presented as the optimal parameter set for the model (ie the estimation of the global minimum $\boldsymbol{\theta}^{**}$ provided by the algorithm). We can also examine the sequence of optimised parameter distributions $\{\varpi_1(\boldsymbol{\theta}), \ldots, \varpi_{k^*}(\boldsymbol{\theta})\}$ resulting from all iterations (Fig. 4.1F) to obtain more information about the convergence and behaviour of the method.

4.3 Application to biological examples

We apply our algorithm to four biological examples of interest. The first three correspond to simulated data from well-known models published in the literature, while in the third example we apply our algorithm to unpublished experimental data of the dynamical response of an inducible genetic promoter constructed for an application in Synthetic Biology.

4.3.1 BPM model of gene-product regulation

The Bliss-Painter-Marr (BPM) model (Bliss *et al.*, 1982) describes the behaviour of a gene-enzyme-product control unit with a negative feedback loop:

$$\begin{aligned} \dot{R} &= \frac{\alpha}{1+P} - \beta R, \\ \dot{E} &= \beta(R - E), \\ \dot{P} &= \beta E - c(t)\frac{P}{1+P}. \end{aligned} \tag{4.10}$$

Here, R, E and P are the concentrations (in arbitrary units) of mRNA, enzyme and product, respectively. The degradation rate of the product has an explicit time dependence, which in this case has the form of a ramp saturation:

$$c(t) = \begin{cases} 5 + 0.2t & 0 \leq t < 50, \\ 15 & t \geq 50. \end{cases}$$

The model represents a gene that codes for an enzyme which in turn catalyses a product that inhibits the transcription of the gene. This self-inhibition can lead to oscillations, which have been shown to occur in the tryptophan operon in *E. coli* (Bliss *et al.*, 1982). Other systems with similar models and dynamics include calcium signalling, the MAPK cascade and immune responses (Höfer, 1999; Kholodenko, 2000; Stark *et al.*, 2007).

We construct a data set from simulations of this model with $\boldsymbol{\theta}_{\text{real}} = [\alpha, \beta] = [240, 0.15]$ and initial conditions $R(0) = E(0) = P(0) = 0$. In this example, the data set $\mathcal{D}$ consists of 10 measurements of $R(t)$ at particular times with added gaussian noise drawn from

k	Min. Error	$\alpha_k^\ddagger$	$\beta_k^\ddagger$	Conv. $\varpi_k(\alpha)$	Conv. $\varpi_k(\beta)$	ϕ_k
1	56.0941	193.7447	0.1304	-	-	-
2	28.2735	246.7510	0.1528	No	No	133.9020
3	27.2083	248.7557	0.1532	No	No	6.8542
4	26.9838	250.3593	0.1536	No	No	0.6532
5	26.6504	251.7189	0.1538	No	No	0.3281
6	26.6504	251.7189	0.1538	No	No	0.1963
7	26.6504	251.7189	0.1538	Yes	Yes	0.0118
8	26.6504	251.7189	0.1538	No	No	0.0131
9 (k^*)	26.6504	251.7189	0.1538	Yes	Yes	1.414×10^{-6}

Table 4.1: Results of the fitting of the BPM model (4.10) with Algorithm 4.1: smallest error of iteration k; the best values $\alpha_k^\ddagger$ and $\beta_k^\ddagger$; whether the distributions have converged; and the difference of the mean errors of the optimised population.

a $\mathcal{N}(0, 15^2)$ distribution (given in Table B.1). The error function $E_{\mathcal{D}}(\boldsymbol{\theta})$ (4.4) corresponds to a non-convex optimisation landscape[1]: a complex rugged surface with many local minima making global optimisation hard (Fig. 4.1A).

We use Algorithm 4.1 to estimate the 'unknown' parameter values from the 'measurements' of R, as illustrated in Sec. 4.2.3 and Fig. 4.1. Feigning ignorance of the true values, we choose a uniform prior distribution with range $[0, 100]$ for both parameters: $\pi_0(\boldsymbol{\theta}) \sim [U(0, 100), U(0, 100)]$. The rest of the running parameters of the algorithm are set to: $J = 500$, $B = 50$, $p_m = 0.95$ and $Tol = 10^{-5}$. Note that the *true* value of α falls outside of the assumed range of our initial prior, while the range of β in our initial prior is two orders of magnitude larger than its true value. This level of uncertainty about parameter values is typical in data fitting for biological models.

Figure 4.1 highlights a key aspect of our algorithm: the local minimisation can lead to local minima outside of the range of the initial prior. Furthermore, our definition of the historical prior ensures that successive iterations can still find solutions within the largest hypercube of optimised solutions in parameter space. In this example, the algorithm moves away from the $U(0, 100)$ prior for α and finds a distribution around 240 (the true value) after three iterations, while in the case of β, the distribution collapses to values around 0.15 after

[1] I thank Markus Owen of the University of Nottingham for suggesting this example.

one iteration. Although the algorithm finds $\boldsymbol{\theta}^{\ddagger}$ after 5 iterations, the algorithm is terminated after 9 iterations, when the posterior distributions are similar (according to the MW test) and the mean errors have also converged (Table 4.1). The estimated parameters for this noisy data set are $\boldsymbol{\theta}^{\ddagger}_{k^*} = [251.7189, 0.1530]$. In fact, the error of the estimated parameter set is lower than that of the real parameters: $E_{\mathcal{D}}(\boldsymbol{\theta}^{\ddagger}) = 26.65 < E_{\mathcal{D}}(\boldsymbol{\theta}_{\text{real}}) = 28.26$, due to the noise introduced in the data. When a data set without noise is used, the algorithm finds the true value of the parameters to 9 significant digits (not shown).

4.3.2 SIR epidemics model

Susceptible-Infected-Recovered (SIR) models are widely used in epidemiology to describe the evolution of an infection in a population (Anderson and May, 1992). In its simplest form, the SIR model has three variables: the susceptible population S, the infected population I and the recovered population R:

$$\begin{aligned} \dot{S} &= \alpha - (\gamma I + d)S, \\ \dot{I} &= (\gamma S - v - d)I, \\ \dot{R} &= vI - dR. \end{aligned} \tag{4.11}$$

The first equation describes the change in the susceptible population, growing with birth rate α and decreasing by the rate of infection γIS and the rate of death dS. The infected population grows by the rate of infection γIS and decreases by the rate of recovery vI and the rate of death dI. The recovered population grows by the rate of recovery vI and decreases by the death rate dR. Here we use the same form of the equations as Toni *et al.* (2009).

The data generated from the model (4.11) (Table B.2) were obtained directly from Toni *et al.* (2009). Hence the original parameter values were not known to us and further we assumed the initial conditions also to be unknown and fitted them as parameters. We used Algorithm 4.1 to estimate α, γ, v, and d and initial conditions S_0, I_0, and R_0. The prior marginal distributions for all parameters were set as $U(0, 100)$. The other parameters were set to: $J = 1000$, $B = 50$, $p_m = 0.95$ and $Tol = 10^{-5}$. The algorithm converged

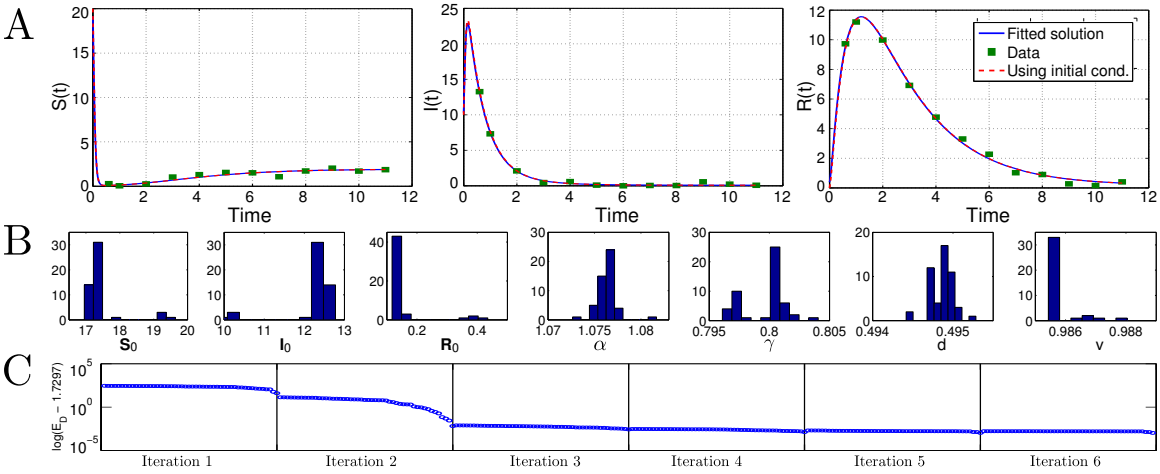

Figure 4.2: Results of fitting SIR model (4.11) to data. **A**: Time courses of the SIR model (4.11). Squares are simulated 'data' points (Table B.2) and bold lines are the model fit with the best parameters $\alpha^{\ddagger} = 1.0726$, $\gamma^{\ddagger} = 0.7964$, $d^{\ddagger} = 0.4945$, and $v^{\ddagger} = 0.9863$ and the best fit initial conditions $S_0^{\ddagger} = 19.1591$, $I_0^{\ddagger} = 10.3016$, and $R_0^{\ddagger} = 0.3861$. Dashed lines use the best fit parameters and the real initial conditions. The minimum error is $E_{\mathcal{D}}(\boldsymbol{\theta}^{\ddagger}) = 1.7297$. **B**: Histogram of the values of the 50 best parameters and initial conditions of the model obtained after convergence at six iterations. **C**: Convergence of the error of the optimised samples at every iteration relative to the final error.

after six iterations. Figure 4.2A shows the prediction of the model (4.11) with the best parameters estimated by our algorithm. The fit is good, with little difference between the curves obtained using the real initial conditions and the ones estimated by our method.

The posterior distributions after six iterations of the algorithm are shown on Fig. 4.2B. The errors obtained after each local minimisation in a decreasing order on each iteration are shown on a semilogarithmic scale in Fig. 4.2C. We can observe how the errors decrease several orders of magnitude over the first three iterations and converge steadily during the last three iterations until $\phi_k \leq Tol$.

4.3.3 An inducible genetic switch from Synthetic Biology

The use of inducible genetic switches is widespread in synthetic biology and bio-engineering as building blocks for more complicated gene circuit architectures. An example is shown schematically in the inset of Fig. 4.3A. This environment-responsive switch is used to control the expression of a target gene G (usually tagged with green

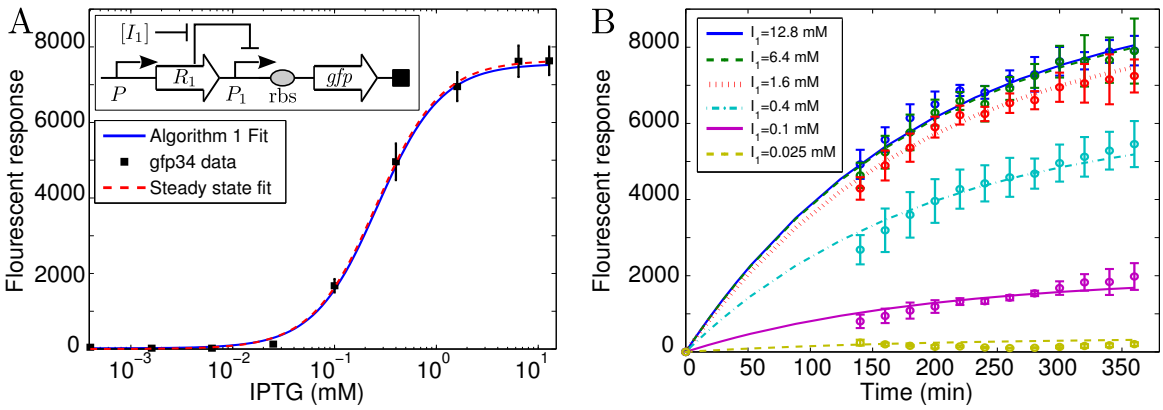

Figure 4.3: Results of fitting genetic switch model (4.12) to data. **A**: *Inset*: An inducible genetic switch consisting of P_1, a negatively regulated environment-responsive promoter. The repressor R_1 promoted by P regulates P_1. The switch is responsive to an exogenous inducer I_1, which binds to R_1 to relieve its repression on P_1 and to turn on the transcription of the downstream target gene, such as a *gfp*. The ribosome binding site (rbs) is used to tune the translation efficiency of the downstream gene. *Plot*: Fluorescent response of the switch with *gfp*-34 to different doses of IPTG (squares). Solution of Eq. (4.12) using the parameters obtained with Algorithm 4.1 (solid line), and stationary solution (dashed line). **B**: Time course of the fluorescent response of the switch with *gfp*-34 to several doses of IPTG (circles) and time-dependent solutions of Eq. (4.12) using the parameters obtained with Algorithm 4.1 (solid lines). Similarly good fits were obtained for responses to I_1 = 0.0063, 0.0016, 0.0004, and 0.0 mM (not shown).

fluorescent protein or *gfp*) through the addition of an exogenous small molecule I_1 (eg isopropyl thiogalactopyranoside or IPTG). The input-output behaviour of this system can be described by the following ordinary differential equation (Alon, 2007; Szallasi *et al.*, 2006):

$$\dot{G} = \alpha k_1 + \frac{k_1 I_1^{n_1}}{K_1^{n_1} + I_1^{n_1}} - dG. \tag{4.12}$$

Here, αk_1 is the basal activity of the promoter P_1 and dG is the linear degradation term. The second term is a Hill function that models the cooperative transcription activation in response to the inducer I_1 with maximum expression rate k_1, Michaelis constant K_1 and Hill coefficient n_1.

The $lacI$–P_{lac} switch has been characterised experimentally in response to different doses of IPTG by Wang (2010); Wang *et al.* (2011). Equation (4.12) can be solved explicitly and one can use nonlinear least squares and the analytical solution to fit data

	Wang (2010)		Algorithm 4.1	
Parameter	*gfp*-30	*gfp*-34	*gfp*-30	*gfp*-34
$\alpha^{\ddagger}$	0.0012 ± 0.027	1.4720×10^{-9}	0.0043	0.0024
$k_1^{\ddagger}$	N/A	N/A	76.1354	63.6650
$n_1^{\ddagger}$	1.3700 ± 0.270	1.3690 ± 0.021	1.4832	1.3879
$K_1^{\ddagger}$	0.2280 ± 0.039	0.2590 ± 0.021	0.2467	0.2641
$d^{\ddagger}$	N/A	N/A	0.0069	0.0052
$k_1^{\ddagger}/d^{\ddagger}$	9456 ± 487	7648 ± 152	10983.34	12163.04

Table 4.2: Parameter values obtained from *gfp*-30 and *gfp*-34 data. In (Wang, 2010), only the steady state solution was used. Hence only the ratio of k_1 and d can be estimated.

at stationarity (ie at long times) and estimate α, n_1, K_1, and the ratio k_1/d. These estimates have been obtained assuming equilibrium ($\dot{G} = 0$) and initial condition $G(0) = 0$ by Wang *et al.* (2011) (Table 4.2).

In fact, the experiments measured time series of the expression of G every 20 minutes from $t = 140$ to 360 min. for different doses of inducer

$$I_1 = 0.0, 3.9 \times 10^{-4}, 1.6 \times 10^{-3}, 6.3 \times 10^{-3}, 2.5 \times 10^{-2}, 0.1, 0.4, 1.6, 6.4, 12.8 \text{ mM},$$

with two different reporters (*gfp*-30 and *gfp*-34, Tables B.4 and B.5). Instead of assuming equilibrium and using only the data for $t > 300$ min as done previously (Wang *et al.*, 2011), we apply Algorithm 4.1 to all the data with the full dynamical equation (4.12) to estimate $\boldsymbol{\theta} = [\alpha, k_1, n_1, K_1, d]$. In this case, we used initial priors $U(0, 1)$ for α and n_1; and $U(0, 20)$ for k_1, K_1 and d. The other parameters were set to: $J = 1000$, $B = 50$, $p_m = 0.95$, and $Tol = 10^{-5}$.

The algorithm converged after five iterations to the parameter values in Table 4.2. The parameter estimates provide good fits to both the time courses (Fig. 4.3B) and to the dose response data (Fig. 4.3A). The values of $K_1^{\ddagger}$ and $n_1^{\ddagger}$ obtained here are similar those obtained in (Wang, 2010) by using only stationary data; this is reassuring since these parameters are related to the dose threshold to half maximal response and to the steepness of the sigmoidal response, both static properties. On the other hand, the values of α and the ratio k_1/d differ to some extent due to the (imperfect) assumption (Wang, 2010) that steady state had been

reached at $t = 300$ min. As Fig. 4.3B shows, G is not at steady state then. Hence the parameter values obtained with our method should give a more faithful representation of the true dynamical response of the switch.

4.3.4 Three-node Repressilator

We test our algorithm with a circuit of genetic regulators, known as the repressilator (Elowitz and Leibler, 2000), which consists of an unidirectional ring of genetic repressors. In one of its simplest forms the system of ODEs that describes the dynamics of the circuit are:

$$\begin{aligned}
\frac{\mathrm{d}m_1}{\mathrm{d}t} &= \alpha_0 + \frac{\alpha}{1+p_3^n} - m_1, \\
\frac{\mathrm{d}p_1}{\mathrm{d}t} &= \beta(m_1 - p_1), \\
\frac{\mathrm{d}m_2}{\mathrm{d}t} &= \alpha_0 + \frac{\alpha}{1+p_1^n} - m_2, \\
\frac{\mathrm{d}p_2}{\mathrm{d}t} &= \beta(m_2 - p_2), \\
\frac{\mathrm{d}m_3}{\mathrm{d}t} &= \alpha_0 + \frac{\alpha}{1+p_2^n} + -m_3, \\
\frac{\mathrm{d}p_3}{\mathrm{d}t} &= \beta(m_3 - p_3).
\end{aligned} \tag{4.13}$$

The variables m_i and p_1 represent the concentrations of mRNA and protein from gene i. The equations for m_1, m_2, and m_3 display a basal level of transcription which is obstructed (by means of a Hill-function) by proteins p_3, p_1, p_2, respectively and have linear decay. The equations for p_i show that translation is proportional to the amount of mRNA and also have linear decay.

The dynamics of model (4.13) are capable of attaining stable steady states, long lived oscillatory transients, and stable oscillations (Elowitz and Leibler, 2000; Strelkowa and Barahona, 2011). The parameters of the model $\boldsymbol{\theta} = [\alpha_0, \alpha, n, \beta]$ are the rates of translation, the Hill-coefficient of the repressor, and the protein translation and decay rate. We use data taken from (Toni and Stumpf, 2010) consisting of measurements of m_1, m_2, and m_3 at several time points with added gaussian noise from simulations using parameters $\alpha_0 = 1$,

$n = 2$, $\beta = 5$, and $\alpha = 1000$, starting from initial conditions $m_1(0) = 0$, $p_1(0) = 2$, $m_2(0) = 0$, $p_2(0) = 1$, $m_3(0) = 0$, and $p_3(0) = 3$ (Table B.3).

To find the value of the parameters using Algorithm 4.1 we set the prior distribution of all the parameters to be $U(0, 500)$ (note that the prior does not include the value of α and contains values much larger than the other parameters' real values), and running parameters $J = 1000$, $B = 50$, $p_m = 0.95$, and $Tol = 10^{-2}$. Figure 4.4A shows all the local minima obtained by our method in different iterations projected into two-parameter scatterplots (the 100 best points of each iteration). As in the BPM example, the local minima align with the landscape of the error function. In the early iterations the algorithm explores regions of the parameter space several orders of magnitude across before focusing on a specific region. Figure 4.4B shows the errors of the local minima found after each iteration. After the first two iterations there was not a significant reduction of the error; until iteration three the algorithm was able to find regions in the parameter space with lower errors. The method found a second "plateau" of the error until iteration 13, then the error decreased steadily until convergence at the end of iteration 20. The minimum error was $E_{\mathcal{D}}(\boldsymbol{\theta}^{\ddagger}) = 484.53$, where $\alpha_0^{\ddagger} = 0.9959$, $n^{\ddagger} = 1.9703$, $\beta^{\ddagger} = 4.7840$, and $\alpha^{\ddagger} = 1,043.90$. Note that $E_{\mathcal{D}}(\boldsymbol{\theta}_{\text{Real}}) = 500.59$, again an artifact of the noise in the data. Figures 4.4C and 4.4D show the final distribution of the parameters and the time-course of the solution with the data. This example illustrates an important aspect of parameter fitting tasks; the algorithm may remain trapped in regions of the parameter space for several iterations, emphasising the need for careful consideration of the choice of priors and running parameters.

4.4 Discussion

We have presented an optimisation algorithm in this chapter that brings together ingredients from Evolutionary Algorithms, local optimisation and Sequential Monte Carlo. The method is particularly useful for determining parameters of ordinary differential equation models from data; however, this approach can also be used in other contexts where an optimisation problem has to be solved on complex landscapes, or when the objective function cannot be written explicitly. The algorithm proceeds by generating a population of solutions through Monte Carlo sampling from a prior distribution and refining those

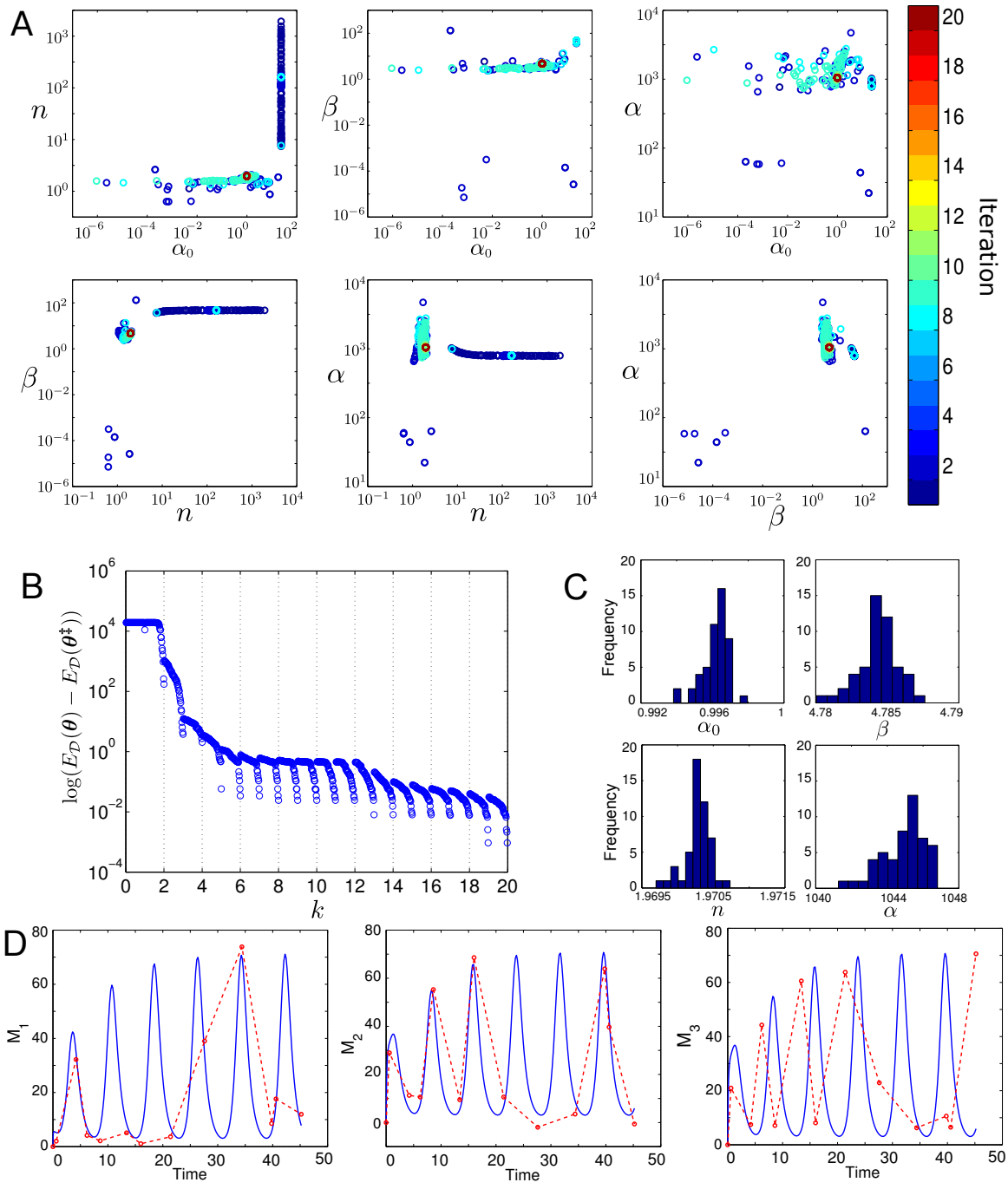
A
B
C
D
Iteration
Frequency
Time
M1
M2
M3

Figure 4.4 *(preceding page)*: Results of fitting repressilator model (4.13) to data. **A**: Scatterplot projections of the 100 best parameter sets after each iteration. Blue markers show minima from early iterations, red markers; from later iterations. **B**: The lowest errors attained after each iteration of the algorithm (on a logarithmic scale, normalised by $E_{\mathcal{D}}(\boldsymbol{\theta}^{\ddagger})$). **C**: Posterior distribution of the parameters at the end of iteration 20. **D**: Time-course of model (4.13) using parameters $\boldsymbol{\theta}^{\ddagger}$ (bold lines, see text) along with the noisy data from Table B.3.

solutions through a combination of local optimisation and culling. A new prior is then created as a mixture of a historical prior (which records the broadest possible range of solutions found) and the distribution of the optimised population. This iterative process induces a strong concentration of the Monte Carlo sampling through local optimisation which accelerates convergence and increases precision, while at the same time the presence of the historical prior allows the possibility that solutions can be found outside of the initial presumed ranges for the parameter values.

The application of the algorithm to ODE models of biological interest has been illustrated and found to perform efficiently. Chapter 7 shows that the algorithm also works well when applied to a larger problem with tens of parameters in the model of stomatal closure. The efficiency of the algorithm hinges on selecting appropriate running parameters and priors. For instance, the number of samples from the prior J should be large enough to allow for significant sampling of the parameter space while small enough to limit the computational cost. We find that simulating $J = 350 - 500$ points in models of up to 10 parameters and keeping the best 15% of the local minima leads to termination within fewer than 20 iterations. In the implementations in this thesis, the Nelder-Mead minimisation is capped at 300 evaluations. These guidelines would result in up to 300,000 evaluations of the objective function per iteration. Therefore the method can become computationally costly if the objective function is expensive to evaluate, eg in stiff models that are difficult to solve numerically. In essence, this algorithm proposes a trade-off: fewer but more costly iterations. It is important to remark that, as with any other optimisation heuristic for non-convex problems, there are no strict guarantees of convergence to the global minimum. Therefore, it is always advisable to run the method with different starting points and different settings with enough sampling points in parameter space to check for consistency of the solutions obtained.

The generation of iterative samples of the parameters draws inspiration from Monte Carlo methods (Sisson *et al.*, 2007; Toni and Stumpf, 2010; Toni *et al.*, 2009) but without pursuing the strict guarantees that the nested structure of the distributions in ABC-SMC provides. This evolutionary approach adopts a highly focused Monte Carlo sampling driven by a sharp local search with culling. Hence this iterative procedure generates samples that only reflect properties of the set of local minima (up to numerical cutoffs) without any focus on the global convergence of the distributions. As noted by Toni *et al.* (2009), the distributions of the parameters (both their sequence and the final distributions) give information about the sensitivity of the parameters: parameters with narrow support will be more sensitive than those with wider support. Future developments of the method will focus on establishing a suitable theoretical framework that facilitates its use in model selection. Broadening the choice of historical priors may be a way of establishing such framework. Currently, no assumptions about the parameter space are made, hence uniform distributions over the support of all the posteriors are used. However, other distributions (eg exponential or log-normal) may be considered as a way to bias the historical prior towards regions of particular interest. Other work will consider the possibility of incorporating a stochastic ranking strategy in the selection of solutions, similar to the one present in the SRES algorithm (Runarsson and Yao, 2000), in order to solve more general constrained optimisation problems with complex feasible regions.

Chapter 5

Models of linear activation cascades: properties and applications

5.1 Introduction

Activation cascades are frequently found in biological signal transduction systems (Heinrich *et al.*, 2002; Marks *et al.*, 2009). Perhaps one of the best studied examples is the *mitogen-activated protein kinase* (MAPK) cascade, which plays a central role in important cellular functions such as regulation of the cell cycle, stress responses and apoptosis (Marks *et al.*, 2009). In general, activation cascades are formed by a set of components (typically proteins) that become sequentially active in response to an external signal (Fig. 5.1). The role of cascades is to relay, amplify, dampen or modulate signals in order to achieve a variety of cellular responses. Activation cascades, particularly the MAPK cascade, have been the subject of numerous studies, experimental and theoretical (Chang and Karin, 2001; Chaves *et al.*, 2004; Feliú *et al.*, 2011; Heinrich *et al.*, 2002; Huang and Ferrell, 1996; Kholodenko, 2000; Tyson *et al.*, 2003; Zhang and Klessig, 2001).

In this chapter, we study ODE models of linear cascades and obtain analytical solutions in terms of the lower incomplete gamma function for the case when inactivation rates are identical and as well as the case when a single protein has a different inactivation rate than the rest. We discuss how these results may be used in parameter fitting and model reduction as an alternative to delay differential equations. The results from this chapter are used to

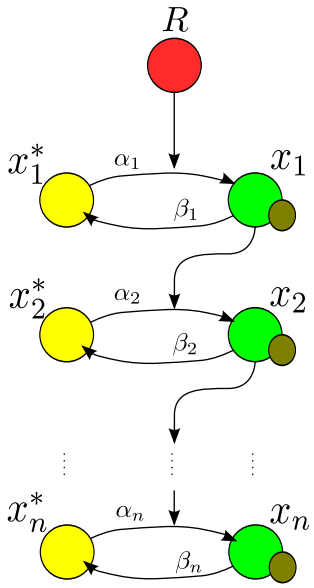

Figure 5.1: A cascade of length n. The nodes in the cascade can either be in an inactive state x_i^*, or active x_i . An external signal $R(t)$ activates the first node. Once a node is active, it can activate the next node, and so on until the end. The activation rates are α_i and the inactivation rates of each x_i are β_i. Image adapted from Heinrich *et al.* (2002).

construct a model of stomatal closure in Chapter 7.

5.2 Linear cascades and their gamma function representation

Consider a cascade of length n subject to an external signal $\hat{R}(t)$. Upon perception of $\hat{R}(t)$, the first inactive component (x_1^*) is transformed into its active state (x_1) which then activates the next inactive component (x_2^*). Sequential activation of x_i^* by x_{i-1} continues until the end of the cascade. The output of the cascade is the active form of the last protein, x_n. In the case of the MAPK cascade, the components are proteins and the activation corresponds to a post-translational modification, ie phosphorylation (Fig. 5.1). However, the formalism below can also be used to describe other sequential biochemical processes with similar functional relationships, for example n-step DNA unwinding (Lucius *et al.*, 2003).

We consider mass-action reactions (without an intermediate complex) for the activation of the proteins. For the activation of x_1 we have

$$\hat{R} + x_1^* \rightarrow x_1 + \hat{R},$$

and for x_i where $i = 2, \ldots, n$ the reaction is

$$x_{i-1} + x_i^* \xrightarrow{\hat{\alpha}_i} x_{i-1} + x_i.$$

We assume that all proteins deactivate spontaneously at a constant rate:

$$x_i \xrightarrow{\beta_i} x_i^*.$$

The ODE model of the reactions above describing the time evolution of the activation cascade is then (Heinrich *et al.*, 2002):

$$\begin{aligned}
\frac{\mathrm{d}x_1}{\mathrm{d}t} &= \hat{R}(t)(T_1 - x_1) - \beta_1 x_1, \\
\frac{\mathrm{d}x_2}{\mathrm{d}t} &= \hat{\alpha}_2 x_1 (T_2 - x_2) - \beta_2 x_2, \\
&\vdots \\
\frac{\mathrm{d}x_n}{\mathrm{d}t} &= \hat{\alpha}_n x_{n-1}(T_n - x_n) - \beta_n x_n,
\end{aligned} \tag{5.1}$$

where $(T_i - x_i) = x_i^*$ is the inactive form of x_i when the total amount of each component is given by $T_i = x_i + x_i^*$. We assume resting initial conditions (ie $x_i(0) = 0$, for all i) and that T_i remains constant (ie no protein production). As shown by Heinrich *et al.* (2002), whenever $T_i \gg x_i$ we have $T_i - x_i \approx T_i$ (the so-called *weakly-activated* case) and we can re-write the system (5.1) as

$$\begin{aligned}
\frac{\mathrm{d}x_1}{\mathrm{d}t} &= R(t) - \beta_1 x_1, \\
\frac{\mathrm{d}x_2}{\mathrm{d}t} &= \alpha_2 x_1 - \beta_2 x_2, \\
&\vdots \\
\frac{\mathrm{d}x_n}{\mathrm{d}t} &= \alpha_n x_{n-1} - \beta_n x_n,
\end{aligned} \tag{5.2}$$

where $R(t) = \hat{R}(t)T_1$ and $\alpha_i = \hat{\alpha}_i T_i$.

The system of equations (5.2) is linear and can be written in vector form:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + R(t)\mathbf{e}_1, \tag{5.3}$$

where $\mathbf{x} = [x_1, \ldots, x_n]^T$, the $n \times n$ rate matrix $\mathbf{A}$ is:

$$\mathbf{A} = \begin{bmatrix} -\beta_1 & & & \\ \alpha_2 & -\beta_2 & & \\ & \ddots & \ddots & \\ & & \alpha_n & -\beta_n \end{bmatrix}, \tag{5.4}$$

and $\mathbf{e}_1 = [1, 0, \ldots, 0]^T$ is the first $n \times 1$ vector of the canonical basis. In general, we use $\mathbf{e}_i$ to denote the i-th canonical vector.

When a cascade receives an integrable input $R(t)$, the Laplace transform (Appendix C.1.1) of the first protein is:

$$\mathscr{L}(x_1) = \frac{\alpha_{(1)}\mathscr{L}(R)}{(\beta_i + s)}.$$

In general, the transform for the n-th protein is

$$\mathscr{L}(x_n) = \frac{\alpha_{(n)}^n \mathscr{L}(R)}{\prod_{i=1}^n (\beta_i + s)}, \tag{5.5}$$

and $\alpha_{(n)}$ is the geometric mean of the activation rates

$$\alpha_{(n)} = \left(\prod_{j=1}^n \alpha_j \right)^{1/n}. \tag{5.6}$$

Note that if $\beta_i \neq \beta_k$ for all i, k then

$$\prod_{i=1}^n (s + \beta_u)^{-1} = \sum_{i=1}^n \frac{\beta_{(-i)}}{s + \beta_i},$$

where

$$\beta_{(-i)} = \prod_{\substack{k=1 \\ k\neq i}}^{n} (\beta_i - \beta_k)^{-1} \in \mathbb{R},$$

we can express equation (5.5) as

$$\mathscr{L}(x_n) = \sum_{i=1}^{n} \frac{\alpha_{(n)}^n \beta_{(-i)} \mathscr{L}(R)}{s + \beta_i}. \tag{5.7}$$

Using the linearity of the inverse of the Laplace transform and its convolution properties we have that the output of the cascade is

$$x_n(t) = \alpha_{(n)}^n \sum_{i=1}^{n} \beta_{(-i)} (R * e^{-t})(\beta_i t), \tag{5.8}$$

where

$$(R * e^{-t})(\beta_i t) = \int_0^{\beta_i t} e^{-(\beta_i t - \tau)} R(\tau) \mathrm{d}\tau = \int_0^{\beta_i t} e^{-\tau t} R(\beta_i t - \tau) \mathrm{d}\tau. \tag{5.9}$$

Example 5.2.1. Consider a cascade of length n with a constant stimulus $R(t) = \alpha_1$, $t \geq 0$, then the output of the last protein in the cascade is given by the sum of exponential functions:

$$x_n(t) = \alpha_{(n)}^n \sum_{i=1}^{n} \beta_{(-i)} \left[1 - e^{-\beta_i t}\right].$$

5.2.1 Optimal cascades

Cells operate with limited resources and they must use them efficiently. Activation cascades as part of the cell-signalling machinery should operate as economically as possible, minimising use of valuable resources such as ATP or amino acids. Recently, Chaves *et al.* (2004) showed that activation cascades in the weakly activated regime (whose dynamics are described by equation (5.3)) are optimal for a given gain (ie achieve maximal amplification) when all the deactivation rates of the proteins are equal and the number of proteins is finite. The gain of a linear cascade is defined as the supremum of equation (5.7):

$$G_n = \frac{1}{l} \frac{\alpha_2 \cdots \alpha_n}{\beta_1 \cdots \beta_n},$$

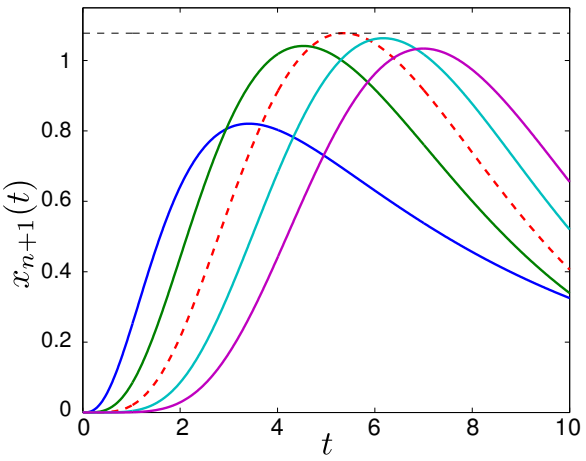

Figure 5.2: Time-course of the leaky integrators ($x_{n+1}(t)$) of cascades of lengths 1 to 5 with fixed gain (Example 5.2.2). The optimal length of the cascade is 3 (dashed red-line), whose maximum value is 1.0779 (indicated by the horizontal dashed line).

where l is the degradation of a protein (x_{n+1}) introduced in the equations at the bottom of the cascade with activation rate equal to 1 (ie a leaky-integrator). As mentioned, given a fixed value of G_n, the cascade provides optimal amplification when $\beta_i = \beta$ for all i (Chaves *et al.*, 2004). This result means that arbitrarily long cascades are not useful for cells that require particular gains from external signals. Furthermore, the degradation of the proteins is

$$\beta = \left(\frac{\alpha_2 \cdots \alpha_n}{l G_n} \right)^{1/n},$$

and the rate-matrix in equation (5.3) becomes

$$\tilde{\mathbf{A}} = \begin{bmatrix} -\beta & & & \\ \alpha_2 & -\beta & & \\ & \ddots & \ddots & \\ & & \alpha_n & -\beta \end{bmatrix}. \tag{5.10}$$

Example 5.2.2. Let $G_n = 7$, $R(t) = 5te^{-2t}$, and $\alpha_i = 1.2$. We examine the time-course of $x_{n+1}(t)$ (the leaky integrator) in cascades of length $n = 1, \ldots, 5$, shown in Fig. 5.2. The optimal cascade for the prescribed G_n has length $n = 3$ (red-dashed line), which means that $\beta = 0.6273$. Longer or shorter cascades exhibit suboptimal amplification.

Below we consider three cases where the input of the cascade can be of particular importance in the modelling process.

Constant stimulus

In an experimental setting, one often wants to study the response of a biological system to a constant stimulus such as constant temperature, light or treatment. In these cases we express the stimulus as:

$$R(t) = \alpha_1 \in \mathbb{R},\ t \geq 0.$$

Then the solution to equation (5.3) with initial condition $\mathbf{x}(0) = \mathbf{0}$ is:

$$\mathbf{x}(t) = \alpha_1 \mathbf{A}^{-1} \left[e^{t\mathbf{A}} - \mathbf{I}_n \right] \mathbf{e}_1, \tag{5.11}$$

where $\mathbf{I}_n$ is the $n \times n$ identity matrix, and $e^{t\mathbf{A}}$; the matrix exponential. If the cascade is optimal (ie $\mathbf{A} = \tilde{\mathbf{A}}$), the Laplace transform of the last protein in (5.5) becomes

$$\mathscr{L}(x_n) = \frac{\alpha_{(n)}^n}{s(s+\beta)^n},$$

whose inverse transform is (see Appendix C.1.3 for detailed calculations):

$$x_n(t) = \left(\frac{\alpha_{(n)}}{\beta} \right)^n \mathrm{P}(n, \beta t), \tag{5.12}$$

where $\mathrm{P}(n, \beta t)$ is the normalised lower incomplete gamma function (defined in equation (C.4)).

Exponentially decreasing stimulus

When the first protein in the cascade is subject to an exponentially decaying stimulus (eg when the signal is a reactive molecule or it becomes metabolised, or the receptors become desensitised)

$$R(t) = \alpha_1 e^{-\lambda t},$$

then the solution to equation (5.3) with initial condition $\mathbf{x}(0) = \mathbf{0}$ is

$$\mathbf{x}(t) = \alpha_1 \left[e^{t\mathbf{A}} - e^{-\lambda t}\mathbf{I}_n\right] \mathbf{A}^{-1} \left[\mathbf{I}_n + \lambda \mathbf{A}^{-1}\right]^{-1} \mathbf{e}_1. \tag{5.13}$$

If we assume again that $\mathbf{A} = \tilde{\mathbf{A}}$, then

$$\mathscr{L}(x_n) = \frac{\alpha_{(n)}}{(s+\lambda)(s+\beta)^n}$$

and the output of the system is given by:

$$x_n(t) = \begin{cases} \left(\frac{\alpha_{(n)}}{\beta-\lambda}\right)^n e^{-\lambda t}\, \mathrm{P}(n, (\beta-\lambda)t) & \text{if } \beta \neq \lambda \\ \\ \frac{1}{\Gamma(n+1)} \left(\alpha_{(n)} t\right)^n e^{-\beta t} & \text{if } \beta = \lambda, \end{cases} \tag{5.14}$$

where $\alpha_{(n)}$ is defined in (5.6) and $\Gamma(n+1)$ is the gamma function. As for the case of constant stimulus, the solution is also given in terms of the lower incomplete gamma function (see Appendix C.1.4 for calculations).

Sinusoidal stimulus

In certain experimental settings one is interested in studying the response of a system to a periodic stimulus, for example in models of circadian rhythms, or periodic stimuli such as day/night cycles (Locke *et al.*, 2005). For that purpose we consider a cascade of length n with a periodic input

$$R(t) = \alpha_1 \left(1 + \sin\left(\omega t\right)\right),$$

that oscillates between 0 and α_1 with frequency $\omega > 0$. From resting initial conditions (ie $\mathbf{x}(0) = \mathbf{0}$), the solution to equation (5.3) is:

$$\mathbf{x}(t) = \alpha_1 \mathbf{V}^{-1} \left[\left(e^{t\mathbf{A}} - \mathbf{I}_n\right) \mathbf{V} - \left(\sin\left(\omega t\right)\mathbf{I}_n + \omega \cos\left(\omega t\right)\mathbf{A}^{-1}\right) + \omega \mathbf{A}^{-1} e^{t\mathbf{A}}\right] \mathbf{A}^{-1}\mathbf{e}_1, \tag{5.15}$$

where $\mathbf{V} = (\mathbf{I}_n + \omega^2 \mathbf{A}^{-2})$. When $\mathbf{A} = \tilde{\mathbf{A}}$ we can obtain the explicit solution for the n-th protein in the cascade (see Appendix C.1.5 for calculations):

$$x_n(t) = \left(\frac{\alpha_{(n)}}{\beta}\right)^n \mathrm{P}(n, \beta t) + \left(\frac{\alpha_{(n)}}{r}\right)^n \left[\sin(\omega t - n\theta) - e^{-\beta t} \sum_{k=0}^{n} \frac{t^k r^k}{k!} \cos((n+k)\theta) \right]$$

where $r = (\beta^2 + \omega^2)^{1/2}$, and $\theta = \arctan\left(\frac{\beta}{\omega}\right)$. This solution consists of the normalised lower-incomplete gamma function added to a function that takes a similar shape to the incomplete gamma function with cosines as coefficients in the polynomials. Note that $\cos((n+k)\theta) = T_{n+k}(\cos(\theta))$ is the $(n+k)$th Chebyshev polynomial evaluated at $\cos(\theta)$. For large t, $x_n(t)$ converges to

$$\breve{x}_n(t) = \left(\frac{\alpha_{(n)}}{\beta}\right)^n \left[1 + \left(\frac{\beta}{r}\right)^n \sin(\omega t - n\theta)\right], \tag{5.16}$$

the behaviour in the long term of $x_n(t)$. The term $\left(\frac{\beta}{r}\right)^n$ in equation (5.16) is useful to characterise the response based on the frequency of the stimulus, and could also be used to establish bounds on the response of a cascade to noisy inputs with variance related to ω (note that $\beta/r < 1 \Rightarrow \breve{x}_n(t) > 0 \ \forall t$).

5.3 Perturbation of a single inactivation rate

We now examine how the output of a weakly-activated (linear) activation cascade is modified when a single protein in the cascade has a different inactivation rate. For instance, Chaves *et al.* (2004) considered a cascade with an auxiliary protein with different inactivation rate (ie the leaky integrator x_{n+1} described in the previous section) at the end of the cascade. We study the effect of such a 'perturbation' and how the effect depends on the position of the component in the cascade. Consider a cascade of n proteins with activation rates α_j and inactivation rates $\beta_j = \beta, \ \forall j \neq i$, and $\beta_i = \beta + \varepsilon$ for a given node i. We can see from the Laplace transform of $x_n(t)$ in equation (5.5) that the position of the protein with degradation β_i in the cascade does not affect the final output of the cascade.

We can also examine the rate matrix of the system (5.3) which in this case becomes

$$\mathbf{A} = \tilde{\mathbf{A}} - \varepsilon\, \mathbf{e}_i \mathbf{e}_i^T \equiv \mathbf{H}_i, \tag{5.17}$$

where $\tilde{\mathbf{A}}$ is given in equation (5.10) and $\mathbf{e}_i$ is the i-th vector of the canonical basis in $\mathbb{R}^n$.

The Jordan decomposition of $\mathbf{H}_i$ is

$$\mathbf{H}_i = \mathbf{Q}_i \mathbf{J} \mathbf{Q}_i^{-1}, \tag{5.18}$$

where $\mathbf{J}$ is the Jordan normal form

$$\mathbf{J} = \begin{bmatrix} -(\beta+\varepsilon) & & & & & \\ & -\beta & 1 & & & \\ & & -\beta & \ddots & & \\ & & & \ddots & 1 & \\ & & & & -\beta & \end{bmatrix}, \tag{5.19}$$

and $\mathbf{Q}_i$ is the matrix with generalised eigenvectors as columns.

Interestingly, a property of this Jordan decomposition is that both $\mathbf{J}$ and the vector $\mathbf{Q}_i^{-1}\mathbf{e}_1$ are independent of the location of the perturbation i. As shown in Example C.2.1, this follows from the following fact: consider $i < h$ (without loss of generality), then rows 1 to $i-1$ and h to n of $\mathbf{Q}_i$ and $\mathbf{Q}_h$ are identical, ie (in Matlab notation) $\mathbf{Q}_i(1:i-1,\ :) = \mathbf{Q}_h(1:i-1,\ :)$ and $\mathbf{Q}_i(h:n,\ :) = \mathbf{Q}_h(h:n,\ :)$. (proof in Appendix C.2.)

5.3.1 Constant stimulus

The constant stimulus solution (5.11) for the perturbed case ($\mathbf{A} = \mathbf{H_i}$) is

$$\mathbf{x}(t) = \alpha_1 \mathbf{H}_i^{-1} \left[e^{t\mathbf{H}_i} - \mathbf{I}_n\right] \mathbf{e}_1 = \alpha_1 \mathbf{Q}_i \mathbf{J}^{-1} \left[e^{t\mathbf{J}} - \mathbf{I}_n\right] \mathbf{Q}_i^{-1} \mathbf{e}_1, \tag{5.20}$$

which follows from (5.18). For two cascades with modified decay rates at i and h ($i < h$), we know from (5.5) that in each case $x_n(t)$ is exactly the same. Furthermore, the properties of $\mathbf{J}$ and $\mathbf{Q}_i$ stated above imply that the vector $\mathbf{J}^{-1}\left[e^{t\mathbf{J}} - \mathbf{I}_n\right]\mathbf{Q}_i^{-1}\mathbf{e}_1$ is independent of the

position of the perturbation, i. Hence the entries of $\mathbf{x}(t)$ are determined only by the matrix $\mathbf{Q}_i$. In both perturbed cascades we have $\mathbf{Q}_i(h:n,:) = \mathbf{Q}_h(h:n,:)$, meaning that the solution for the last $n-h+1$ proteins is the same in both cascades. (Similarly, the $i-1$ first components of the solution (5.20) are identical, but that is obvious).

Example 5.3.1. Consider two cascades of length $n = 6$ with constant stimulus and activation rates $\alpha_1 = \cdots = \alpha_6 = 3$ and degradation rate $\beta = 2$ for all proteins except for a perturbation $\varepsilon = 0.5$ on the third and fifth proteins of each cascade, respectively. The corresponding rate matrices are:

$$\mathbf{H}_3 = \begin{bmatrix} -2 & & & & & \\ 3 & -2 & & & & \\ & 3 & -2.5 & & & \\ & & 3 & -2 & & \\ & & & 3 & -2 & \\ & & & & 3 & -2 \end{bmatrix} \quad \mathbf{H}_5 = \begin{bmatrix} -2 & & & & & \\ 3 & -2 & & & & \\ & 3 & -2 & & & \\ & & 3 & -2 & & \\ & & & 3 & -2.5 & \\ & & & & 3 & -2 \end{bmatrix}.$$

The Jordan form for both cascades is:

$$\mathbf{J} = \begin{bmatrix} -2.5 & & & & & \\ & -2 & 1 & & & \\ & & -2 & 1 & & \\ & & & -2 & 1 & \\ & & & & -2 & 1 \\ & & & & & -2 \end{bmatrix}$$

and the corresponding generalised eigenvector matrices are:

$$\mathbf{Q}_3 = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{3} & \mathbf{0} \\ 36 & 0 & 0 & 0 & 18 & -36 \\ -216 & 0 & 0 & 54 & -108 & 216 \\ \mathbf{1296} & \mathbf{0} & \mathbf{162} & \mathbf{-324} & \mathbf{648} & \mathbf{-1296} \\ \mathbf{-7776} & \mathbf{486} & \mathbf{-972} & \mathbf{1944} & \mathbf{-3888} & \mathbf{7776} \end{bmatrix}, \tag{5.21}$$

$$\mathbf{Q}_5 = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{3} & \mathbf{0} \\ 0 & 0 & 0 & 9 & 0 & 0 \\ 0 & 0 & 27 & 0 & 0 & 0 \\ \mathbf{1296} & \mathbf{0} & \mathbf{162} & \mathbf{-324} & \mathbf{648} & \mathbf{-1296} \\ \mathbf{-7776} & \mathbf{486} & \mathbf{-972} & \mathbf{1944} & \mathbf{-3888} & \mathbf{7776} \end{bmatrix}. \tag{5.22}$$

As explained above, rows 1-2 and 5-6 of $\mathbf{Q}_3$ and $\mathbf{Q}_5$ (in bold) are the same. In addition,

$$\mathbf{Q}_3^{-1}\mathbf{e}_1 = \mathbf{Q}_5^{-1}\mathbf{e}_1 = [1\ 0\ 0\ 0\ 0\ 1]^{\mathrm{T}}.$$

Since the rows of of $\mathbf{Q}_3$ and $\mathbf{Q}_5$ are the same below the second perturbation, then the values of $x_5(t)$ and $x_6(t)$ are equal in both cascades. Figure 5.3 shows the time course of the proteins in both cascades: $x_1(t)$, $x_2(t)$, $x_5(t)$, and $x_6(t)$ (solid lines) are the same in both cascades, while $x_3(t)$ and $x_4(t)$, the proteins "sandwiched" between the perturbations (dashed lines), are not. We discuss an application of this result in Section 5.4.2.

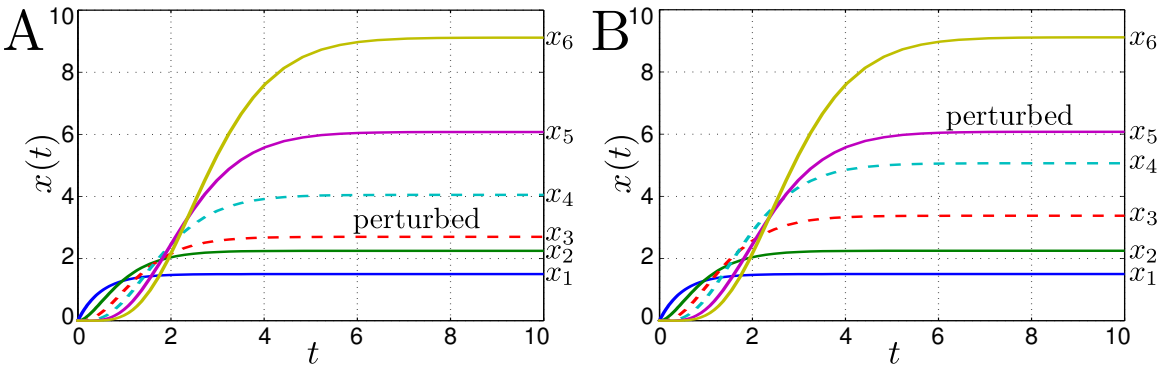

Figure 5.3: Time course solutions of two ε-perturbed cascades from Example 5.3.1. **A**: Cascade with a perturbation in the degradation of the third protein. **B**: Cascade with a perturbation in the fifth protein. The activity of the proteins of both cascades is the same for nodes above and below the perturbations (continuous lines), but is different in the nodes between the perturbations (dashed lines).

5.3.2 Exponentially decreasing stimulus

Just as in the previous section, the solution (5.13) for the exponential stimulus in the perturbed case ($\mathbf{A} = \mathbf{H_i}$) can be rewritten as:

$$\mathbf{x}(t) = \alpha_1 \mathbf{Q}_i \left[e^{t\mathbf{J}} - e^{-\lambda t}\mathbf{I}_n\right] \mathbf{J}^{-1} \left[\mathbf{I}_n + \lambda \mathbf{J}^{-1}\right]^{-1} \mathbf{Q}_i^{-1} \mathbf{e}_1, \tag{5.23}$$

and, again, the same argument follows to conclude that the last $n - h + 1$ components of the solution (5.23) are the same for two cascades modified at positions i and h ($h > i$).

Example 5.3.2. Consider the same cascades as in Example 5.3.1 with $R(t) = \alpha_1 e^{-\lambda t}$ and $\lambda = 1$. Figure 5.4 shows the time behaviour of the proteins in the two cascades. As in the previous example, the proteins above and below the perturbations are unchanged.

5.4 Applications

5.4.1 Model simplification and parameter fitting

The expressions for the output of the cascade $x_n(t)$ in terms of incomplete gamma functions can be useful to fit activation data to a reduced number of parameters. Rather than

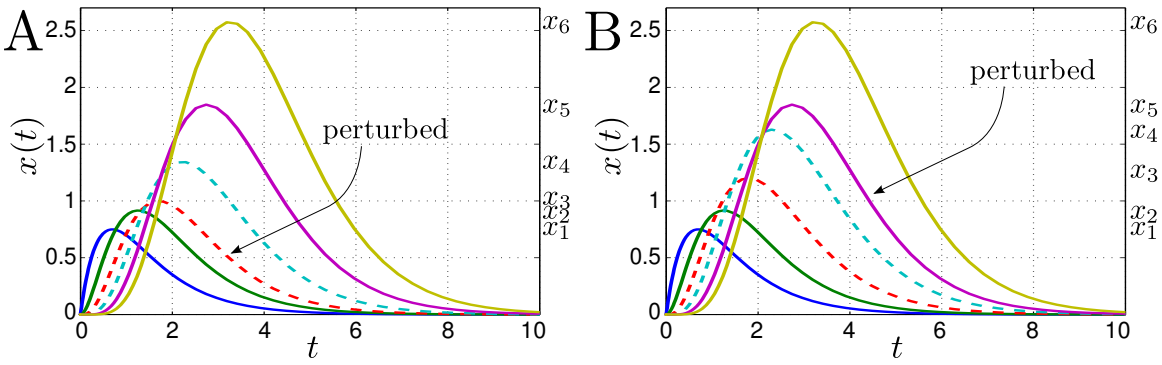

Figure 5.4: Time course solutions of two ε-perturbed cascades with decaying stimulus from Example 5.3.2. **A**: Cascade with a perturbation in the third protein. **B**: Cascade with a perturbation in the fifth protein. On the right of each plot, the label of each protein is placed at the level where its solution peaks. The activity of the proteins is the same for nodes above and below the perturbations (continuous lines), but different in the nodes between the perturbations (dashed lines).

fitting the observed output of the cascade to $n + 2$ parameters for an entire module with n components, the approximate expression with the lower incomplete gamma function contains at most four parameters: $\alpha_{(n)}$, β, n (and in the case of exponential decay, λ). In this approach, (shown graphically in Fig. 5.5A) the length of the cascade n is turned into a fitting parameter, similarly to what is done with Hill coefficients. Indeed, the fitted value does not need to be an integer because the lower incomplete gamma function $\mathrm{P}(n, t)$ is defined for any positive real number in its first argument (Abramowitz and Stegun, 1964).

Example 5.4.1. Consider two cascades of length $n = 5$ with parameters $\alpha_1 = 3$, $\alpha_i = 4$ for $i = 2, \ldots, 5$ (so $\alpha_{(n)} = 3.776$), and $\beta = 3$. One cascade is subject to a constant stimulus $R(t) = \alpha_1$ and the other to an exponentially decaying input $R(t) = \alpha_1 e^{-\lambda t}$ with $\lambda = 1$.

After solving numerically the ODE models of the linear cascade (5.3) for both inputs (dashed lines in Fig. 5.5B), we sample the output $x_5(t)$ at times $t = \{0, 1, \ldots, 10\}$ and we add random noise from a distribution $\mathcal{N}(0, 0.05^2)$ to generate our 'observed data' (squares in Fig. 5.5B).

We fit the gamma function expressions (5.12) and (5.14) to the 'data' using the method introduced in Chapter 4 (Appendix C.4). The bold lines in Fig. 5.5B show the fits to both cascade output data. The fits to the noisy data are good and the estimated values are close to the 'true' ones: in the case of the constant stimulus cascade, the fitted values

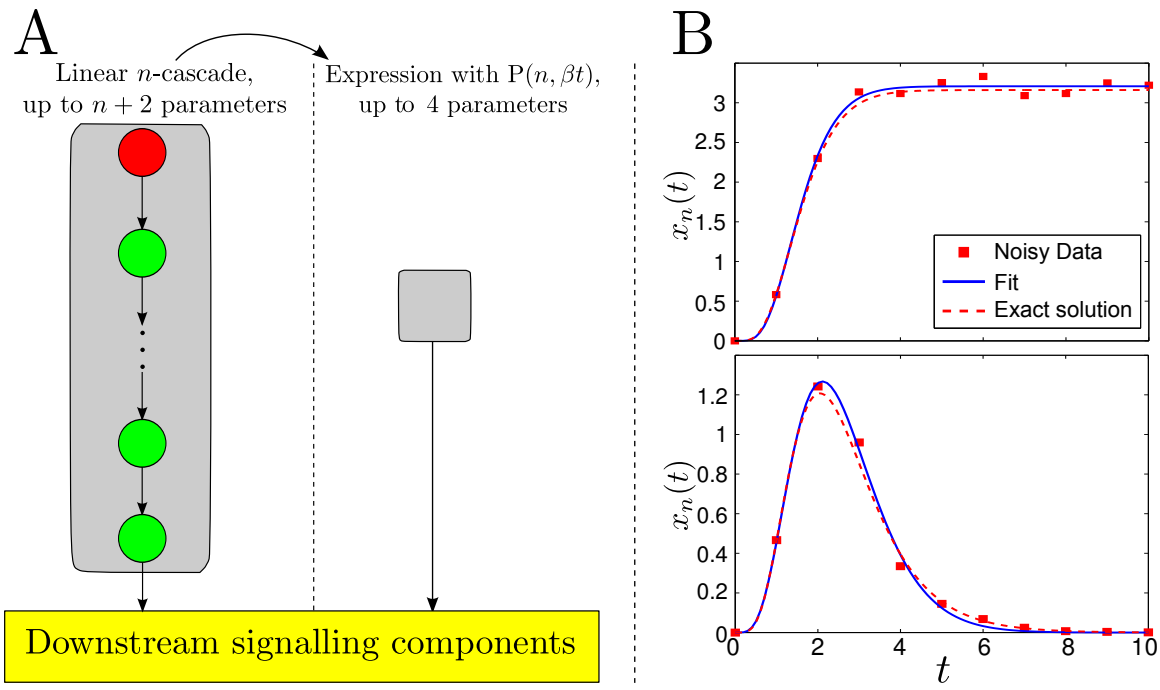

Figure 5.5: **A**: Schematic of a signal transduction model with a cascade. Given a linear cascade of arbitrary length, the red node at the top is the stimulus, and green nodes are the components of the cascade. The last node of the cascade transmits the signal to downstream components of the pathway. The model of the cascade has up to $n + 2$ parameters: $\alpha_1, \ldots, \alpha_n$, β, and if the stimulus decays, λ. Then the cascade is condensed into an expression with an incomplete gamma function that sends the exact same signal as the cascade in the left panel directly to the rest of the network. The new expression has parameters $\alpha_{(n)}$, β, n, and if the stimulus decays, λ. **B**: Examples of the time-course of two cascades with constant (top) and exponentially decaying stimulus (bottom). The dashed lines indicate the solutions to the corresponding systems of ODEs, squares are noisy data generated from the models, and bold lines are fits to the data using the incomplete gamma function expressions (see Example 5.4.1).

are $\alpha_{(n)} \approx 4.068$, $\beta \approx 3.281$, and $n \approx 5.418$; in the case of the exponentially decaying stimulus, the estimated values are $\alpha_{(n)} \approx 3.317$, $\beta \approx 2.177$, $n \approx 4.6$, and $\lambda \approx 2.177$.

5.4.2 Cascade equation reordering

The results presented in Sec. 5.2 and 5.3 allow us to reshuffle equations of cascade models where perturbations are known to occur. In particular, the equations of all proteins with the same inactivation rates can be grouped together upstream in the cascade so that they can be replaced with the incomplete gamma function expression, while the equations of the perturbed proteins are placed downstream and take the gamma function as an input.

This process of reordering the cascade, which is schematically represented in Fig. 5.6A, can be used to reduce the ODE model for the cascade without altering the dynamics or the timescales. Suppose we have an ε-perturbed cascade of $n+1$ proteins that we have reordered so that the first n proteins have inactivation rate β and the $(n+1)$-th protein has rate $\beta + \varepsilon$.

Constant input

Consider first a constant input $R(t) = \alpha_1$. Using (5.12), we write the dynamics of the output of the cascade as

$$\frac{\mathrm{d}x_{n+1}}{\mathrm{d}t} = \alpha_{n+1}\left(\frac{\alpha_{(n)}}{\beta}\right)^n \mathrm{P}(n, \beta t) - (\beta + \varepsilon)x_{n+1}, \tag{5.24}$$

or, more conveniently, its Laplace transform:

$$\mathscr{L}(x_{n+1}) = \frac{\alpha_{(n)}}{s(s+\beta)^{n+1}(s+\beta+\varepsilon)}.$$

This equation can be solved analytically (see C.3.1):

$$x_{n+1}(t) = \frac{\alpha_{n+1}}{\beta+\varepsilon}\left(\frac{\alpha_{(n)}}{\beta}\right)^n \left(1 - e^{-\beta t}\left[\left(\frac{-\beta}{\varepsilon}\right)^n e^{-\varepsilon t} + \sum_{k=0}^{n-1} \frac{(\varepsilon^{n-k} - (-\beta)^{n-k})(\beta t)^k}{\varepsilon^{n-k} k!}\right]\right), \tag{5.25}$$

where we have taken the initial condition $x_{n+1}(0) = 0$.

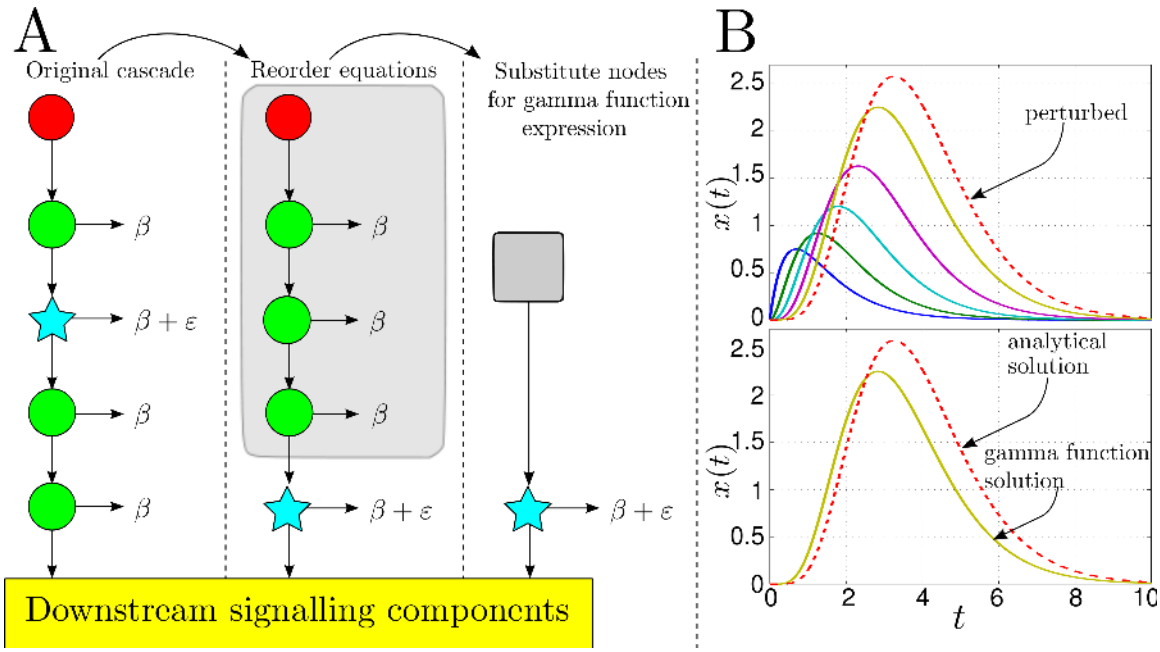

Figure 5.6: Example of cascade reordering and substitution. **A**: A linear ε-perturbed cascade model of length 4, the input (red node) can either be constant or decaying. Green circle nodes are proteins whose inactivation rates are all β, the blue star node has inactivation rate $\beta + \varepsilon$. Downstream of the cascade lie other components of the signalling pathway. Reordering of the equations: the protein with the perturbed inactivation has been moved the bottom of the cascade. Both this cascade and the one on the left have the same output. The first three equations in the reordered cascade are substituted for an incomplete gamma function. **B**: Numerical example of cascade reordering. Top: the equation of the perturbed protein is placed at the bottom of the cascade (the time-course of the untouched cascade is shown in Fig. 5.4A). Bottom: the solution to the module of unperturbed proteins is given by equation (5.14); the solution of the perturbed protein at the bottom of the cascade is given by equation (5.29) (Example 5.4.2).

Exponentially decaying input

Consider an exponentially decaying input $R(t) = \alpha_1 e^{-\lambda t}$. If $\beta \neq \lambda$ the ODE for the $(n+1)$-th protein is

$$\frac{\mathrm{d}x_{n+1}}{\mathrm{d}t} = \alpha_{n+1}\left(\frac{\alpha_{(n)}}{\beta-\lambda}\right)^n e^{-\lambda t}\,\mathrm{P}(n,(\beta-\lambda)t) - (\beta+\varepsilon)x_{n+1}. \tag{5.26}$$

When $\beta = \lambda$ we have:

$$\frac{\mathrm{d}x_{n+1}}{\mathrm{d}t} = \alpha_{n+1}\frac{\left(\alpha_{(n)}t\right)^n}{\Gamma(n+1)}\,e^{-\beta t} - (\beta+\varepsilon)x_{n+1}. \tag{5.27}$$

In both cases the Laplace transform of $x_{n+1}(t)$ is

$$\mathscr{L}(x_{n+1}) = \frac{\alpha_{(n)}^n}{(s+\lambda)(s+\beta)^{n-1}(s+\beta+\varepsilon)}. \tag{5.28}$$

When the initial condition is $x_{n+1}(0) = 0$, the analytical solution for equation (5.26) is (see C.3.2):

$$x_{n+1}(t) = \frac{\alpha_{n+1}}{\beta-\lambda+\varepsilon}\left(\frac{\alpha_{(n)}}{\beta-\lambda}\right)^n \left[e^{-\lambda t} + \frac{e^{-(\beta+\varepsilon)t}}{\varepsilon^n} - \right.$$
$$\left. - e^{-\beta t}\sum_{k=0}^{n-1}\frac{\left(\varepsilon^{n-k} - (\lambda-\beta)^{n-k}\right)(\beta-\lambda)^k t^k}{\varepsilon^{n-k}k!}\right]. \tag{5.29}$$

When $\lambda = \beta$ the solution is

$$x_{n+1}(t) = \left(\frac{\alpha_{(n+1)}}{\varepsilon}\right)^{n+1} e^{-\beta t}\left[\varepsilon^n \sum_{k=0}^{n}\frac{(-1)^k t^{n-k}}{\varepsilon^k (n-k)!} + (-1)^{n+1}e^{-\varepsilon t}\right]. \tag{5.30}$$

Example 5.4.2. Consider the $n = 6$ cascade from Example 5.3.2, where the degradation rate of the third protein is ε-perturbed (time-course shown in Fig. 5.3A). The equations of the system can be reordered without affecting its final output so that the equation of the perturbed protein is at the bottom of the cascade (Fig. 5.6B, top). The output of the first 5 equations in the reordered cascade is then given by the incomplete gamma function expression (5.14) and the analytical solution of the perturbed protein (which is now the

output of the cascade) is given by Eq. (5.29) (Fig. 5.6B, bottom).

5.4.3 Delay differential equation models for activation cascades

Experimental observations in signalling cascades are typically concerned with the effect of the cascade on the output characterised in terms of the amplification, distortion and delay introduced in the output downstream. Within the framework of ODE models, the interactions between the variables occur instantaneously. Hence, if the response to a stimulus occurs with delay, one must incorporate further intermediate variables that were not considered in the original ODE model, corresponding to unmeasured, hidden processes that take time to complete (Stark *et al.*, 2007). This process can lead to large models with many unobservable variables and large numbers of parameters or to the introduction of abstract variables to model unknown processes that may contribute to the observed delays (Bar-Or *et al.*, 2000; Höfer *et al.*, 2002). Alternatively, modellers often use delay differential equations (DDEs) to account for the lag between an event and its effect parsimoniously (Bernard *et al.*, 2006; Colijn and Mackey, 2005; Monk, 2003). In a DDE, the activity of a variable depends on the state of the system in the past:

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{f}(\mathbf{x}(t-\tau)),$$

where the parameter $\tau \geq 0$ is the delay. Linear systems of delay differential equations can be solved analytically using infinite series involving the Lambert function (Bellman *et al.*, 1963; Yi and Ulsoy, 2006), but such solutions are often impractical to use. Our results indicate that simple delays can be modelled with linear activation cascades leading to concise ODE descriptions in terms of the lower incomplete gamma function and without relying on DDEs, as shown in Fig. 5.7A.

Example 5.4.3. Consider a system with a delay, which we model with the following linear DDE:

$$\begin{aligned}\frac{\mathrm{d}\hat{p}_1}{\mathrm{d}t} &= \hat{\alpha} - \hat{\beta}\,\hat{p}_1,\\ \frac{\mathrm{d}\hat{p}_2}{\mathrm{d}t} &= \hat{\alpha}\,\hat{p}_1(t-\tau) - \hat{\beta}\,\hat{p}_2.\end{aligned} \tag{5.31}$$

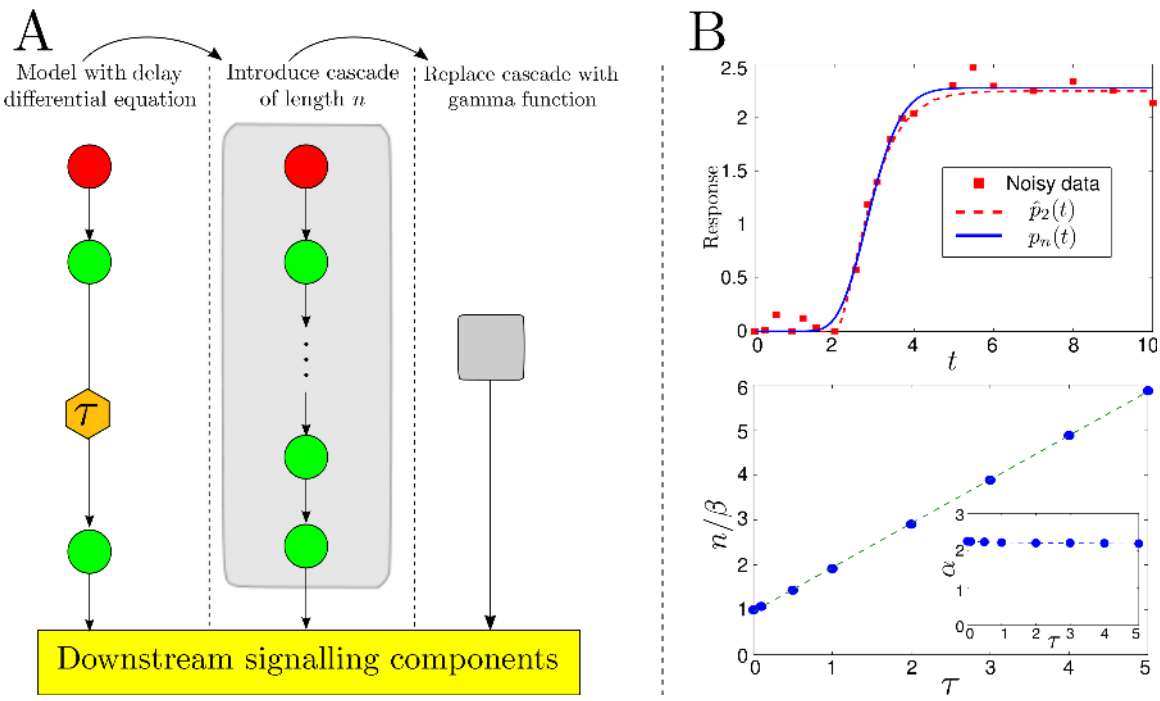

Figure 5.7: **A**: Example of the use of linear activation cascades to replace delay differential equations. A signal node (red node) activates a node in a signalling pathway. The bottom node responds with a delay τ. The delay in the equation is removed and substituted with a linear cascade of length n. The entire cascade is replaced by a lower incomplete gamma function. **B**: Top: The dashed line is the solution to the DDE (5.31) from Example 5.4.3, the squares are points taken from the solution with added random noise. The continuous line is the approximation using a lower incomplete gamma function. Bottom: the relationship between the ratio n/β and τ is linear (dashed-line) with slope 0.977 and intercept 0.962. Inset: α almost does not vary with τ (see text).

Figure 5.7B (top plot) shows the simulated time course of $\hat{p}_2(t)$ (red dashed line) when $\hat{\alpha} = 2$, $\hat{\beta} = 3$, and $\tau = 2$ with initial conditions $\hat{p}_1(0) = \hat{p}_2(0) = 0$. To generate our 'observed data,' we sample $\hat{p}_2$ at various time points and add observational random noise from a distribution $\mathcal{N}(0, 0.05^2)$.

We can fit this noisy data to a linear cascade of length n under constant input with parameters $\alpha_{(n)}$ and β:

$$p_n(t) = \left(\frac{\alpha_{(n)}}{\beta}\right)^n \mathrm{P}(n, \beta t) \approx \hat{p}_2(t), \tag{5.32}$$

and estimate the corresponding parameters. The solid line in the top plot of Fig. 5.7B shows the best fit of the data to a linear cascade, as obtained with the parameter fitting algorithm introduced in Chapter 4 (Appendix C.4). The estimated parameter values are $\alpha_{(n)} \approx 2.27$, $\beta \approx 7.53$, and $n \approx 22.1072$.

We also explore the connection between the parameters of the DDE and the best fitted (linear) activation cascade model, in particular as a function of the delay τ. We simulate the DDE (5.31) with parameters $\hat{\alpha} = 2$ and $\hat{\beta} = 3$) for different values of the delay $\tau \in [0, 5]$ and collect data from these models as above, but without adding random noise. The dependence of the fitted parameters and τ is shown in Fig. 5.7B (bottom plot): α remains relatively constant (decreases minimally) while the ratio n/β grows almost linearly with τ: $n/\beta = 0.962 + 0.977\tau$. In fact, β increases linearly (slope ≈ 4.11) and n grows almost quadratically with τ (exponent ≈ 1.88). When the delay τ in the 'data' is increased, the length of the fitted cascade (n) increases while the time scale of the gamma function ($1/\beta$) decreases in consortium. If one attempts to fit the data allowing only the length n as a fitting parameter, the fit is not successful, thus underscoring the importance of the time scale β in the approximation. Indeed, the original time delay τ in the DDE is approximated in the linear cascade by $(n/\beta - 1)$, ie the accumulated time needed to traverse n sequential steps with duration $1/\beta$.

In Chapter 7 we use the approach described in this section to introduce delays in the antioxidant responses of guard cells to abscisic acid and ethylene stimuli during stomatal closure.

5.5 Discussion

In this chapter the classic model of activation cascades in the weakly activated limit (Heinrich *et al.*, 2002) has been examined. We have considered the important case where all inactivation rates of the components of the cascade are identical, which has been shown to provide optimal amplification (Chaves *et al.*, 2004). These results show that the output of these cascades can be represented exactly by lower incomplete gamma functions. We also show that the position of a protein in the cascade does not affect the final output. These results allow the reduction of the number of equations and parameters in ODE models without affecting the dynamics or the timescales of the system. The results show that in some cases incomplete gamma functions can be used to approximate delay differential equations. Beyond its application to enzymatic activation cascades, similar mathematical models of cascades could be helpful for the parametrisation and modelling of multi-step transcriptional processes, an area of active research in Systems and Synthetic Biology (Hooshangi *et al.*, 2005; Lucius *et al.*, 2003; Stricker *et al.*, 2008; Wang *et al.*, 2011). These results also give an example of how reducing nonlinear models of ODEs is not trivial. Some methods reduce network models (or modules) based on the topology, effectively finding a minimal kernel that preserves some aspects of the dynamics (Kim *et al.*, 2011). By only considering the topology (and the signs) of the system such methods cannot be guaranteed to preserve timescales or behaviour (Ingram *et al.*, 2006), and are best suited for boolean models. As Chapter 7 shows, timescales can be crucially linked to the behaviour of a model. Reducing models of differential equations remains a challenging and active area of research (Conzelmann *et al.*, 2004; Prajna and Sandberg, 2005; Siahaan, 2008). Future work will focus on the characterisation of the output of cascades with negative (nonlinear) feedback (eg see equation (C.50) for a preliminary result), and studying the case in which the degradation of the proteins (β_i) are random variables with the same mean and testing the conditions for optimal amplification given desired expected gains.

Chapter 6

Experimental results

Guard cells were introduced in Chapter 2 along with a description of the way in which plant hormones abscisic acid (ABA) and ethylene induce stomatal closure. The intriguing observations made by Tanaka *et al.* (2005) and Desikan *et al.* (2006), where stomata treated with ABA and ethylene did not achieve full closure have been described in Sec. 2.5. Experiments are performed in this chapter to reproduce such results and to obtain more information about why this unexpected behaviour occurs. The gathering of data in this chapter has been motivated by the development of the models of stomatal closure presented in Chapter 7, where the data are used to fit the model parameters.

Most data about stomatal closure available in the literature show the response of some component of the signalling network, typically aperture, within one to two hours of the stimulus (Bright *et al.*, 2006; Chen *et al.*, 2004; García-Mata and Lamattina, 2007; Li *et al.*, 2000). In addition, the data available are fragmented and sometimes not consistent with each other, for example, with different species such as *Arabidopsis thaliana*, *Vicia faba* (in which ethylene actually mediates stomatal opening (Levitt *et al.*, 1987)) or tomato, laboratory conditions, plant age, cell type (guard cells, epidermal peels, guard cell protoplasts may all have different properties), and methods of measuring stomatal aperture (area or opening width). Detailed temporal analyses of ABA or ethylene-induced signalling responses in guard cells have not yet been made. Moreover, it has often been assumed that ROS production in guard cells only occurs in a burst and therefore, measurements in guard cells are conventionally made over a short period, up to 5 min, after treatment (Pei

and Kuchitsu, 2005; Zhang *et al.*, 2001b). To achieve a better understanding of the signalling processes leading to stomatal closure, it is imperative to know the behaviour of the components of the pathway in time. As larger signalling systems with multiple stimuli become subject of more studies, the fundamental need for time course data is increasingly evident. This study makes a first advance to track the temporal responses of ABA and ethylene in guard cells under single and combined stimuli; to that end, we obtain measurements of ROS production and stomatal aperture in *Arabidopsis* guard cells. The experiments follow the state of the system upon the arrival of the signal for 60 minutes, a time-frame sufficient to study stomatal closure in *Arabidopsis.*

The choice to measure stomatal aperture is clear since it is the phenomenon that is the focus of this study. The choice to measure the concentration of ROS at this point is because it is (with current knowledge) where the pathways of ethylene and ABA in guard cells first meet; additionally, (as explained in Chapter 2) ROS are an important component of the guard cell signalling pathway. It is important to know whether the interactions responsible for the lack of closure happen at this point or further down the pathway. Furthermore, as mentioned above, the time-course response of ROS to either ABA or ethylene stimuli is only known for early time-points; however, the modelling in Chapter 7 indicates that ROS measurements are necessary for longer times, which is why the time-course of ROS here is extended to 60 minutes.

The experiments in this chapter were performed by Mariano Beguerisse and Alessandro Lizzul (stomatal assays), and Mercedes Hernández Gómez (ROS fluorescence assays).

6.1 Materials and methods

6.1.1 Plant material

Arabidopsis thaliana seeds of Columbia ecotype (Col-0) were sown on Levington F2+S (Avoncrop, Bristol, UK) soil and grown under constant conditions in a growth chamber (Sanyo Gallenkamp, Loughborough, UK) with a 10 hour light-period, light intensity of 100-150μ E/m^2/s, temperature of 22°C and 70% relative humidity. After 7 days, seedlings were transplanted individually into new pots. Leaves of 4-6 weeks old plants, which had

not yet formed flower bolts, were harvested for the aim of these experiments.

6.1.2 Stomatal assays

Stomatal aperture bioassays were performed on 4 week old leaves, as described by Desikan *et al.* (2006). Leaves from 4-6 week-old plants with their petioles intact were then cut from the plants using scissors. Excised leaves were floated for 3 hours inside the growth chamber in Petri dishes with MES/KCl buffer (5mM KCl, 50 μM $CaCl_2$, 10mM MES, buffered to pH 6.15 with KOH) to open their stomata. After the initial treatment in buffer the leaves were exposed to doses of 1 μM, 10 μM, and 50 μM ABA (2-cis, 4-trans abscisic acid 98%, synthetic, Aldrich), 1 μM, 10 μM, and 100μM ACC (1-aminocyclopropane-carboxylic acid hydro-chloride, Sigma), or a combination of 10 μM ABA with 10 μM ACC, using ethanol as a control. The treatments were left in the growth chamber for 15, 30, 45 and 60 minutes. Two leaves were blended in water for 1-2 minutes and epidermal fragments collected on a 100 μm nylon mesh (Spectra-Mesh, BDH-Merck, Nottingham, UK) and transferred to a microscope slide. Measurements of individual stomatal aperture were conducted using a Leica DME light microscope, connected to a Leica DFC290 camera imaging system (Leica, Milton Keynes, UK). Leica QWinV3 software (Leica QWIN software, Leica, Milton Keynes, UK) was used to measure the apertures. Each data point collected is the mean of three experiments where $n = 30$ for each experiments (ie 90 individual stomatal measurements).

6.1.3 ROS fluorescence assays

Stomatal H_2O_2 concentration was measured as described by Desikan *et al.* (2006). Leaves from 4-6 week old plants (2 leaves per condition) were blended in deionised water and epidermal fragments were collected with a 40 μm sterile cell strainer (Fisher Scientific). Epidermal fragments were incubated in Petri dishes containing MES/KCl buffer (5mM KCl, 50 μM $CaCl_2$, 10mM MES buffered to pH 6.15 with KOH) inside the growth chamber for 3 hours. Epidermal fragments were collected and equally distributed into Petri dishes loaded with ethanol (as control), 10 μM ABA (2-cis, 4-trans abscisic acid 98%, synthetic, Aldrich), 10 μM the ethylene precursor ACC (1-aminocyclopropane-carboxylic

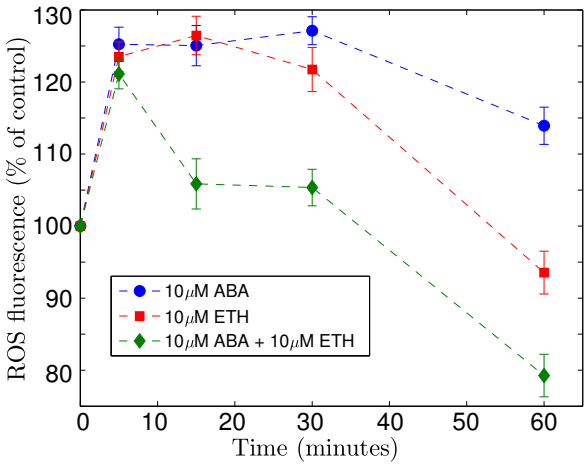

Figure 6.1: ROS fluorescence measurements. Experimental measurements of ROS in guard cells as percentage of control. in response to 10 μM ABA (blue circles), 10 μM ethylene (red squares), and 10 μM ABA + 10 μM ethylene (green diamonds). Error bars show the standard error of the mean ($n = 30 \times 3$).

acid hydro-chloride, Sigma), and a combination of 10 μM ABA and 10 μM ACC. The treatments, which had durations of 5, 15, 30, and 60 minutes, were performed. Following treatment for the appropriate time, fragments were incubated with 50μM H_2DCF-DA (2, 7 -dichlorodihydrofluorescein diacetate, Invitrogen) for 10 minutes for H_2O_2 detection. After a washing step in MES/KCl buffer for 20 minutes, epidermal fragments were placed onto a slide and observed under a microscope. All steps were carried out under dark conditions, as the dye is light-sensitive. H_2O_2 was visualised with a fluorescence microscope (Axioskop2 plus, Carl Zeiss Ltd., UK) with Zeiss filter set 3 (excitation light filter: 450-490nm, emission light filter: 515-565nm). Photographic images were captured with Axiovision software v3.1 (Carl Zeiss Vision GmbH, UK). The images were processed and fluorescence intensities (as mean of the pixel intensities) were measured with ImageJ software (Abramoff *et al.*, 2004).

6.2 ROS production

We applied single and combined doses of 10 μM ABA and 10 μM ethylene (ie ABA, ethylene, and ABA+ethylene) to *Arabidopsis thaliana* guard cells and measured the resulting ROS concentration in fluorescence units. The level of ROS production relative

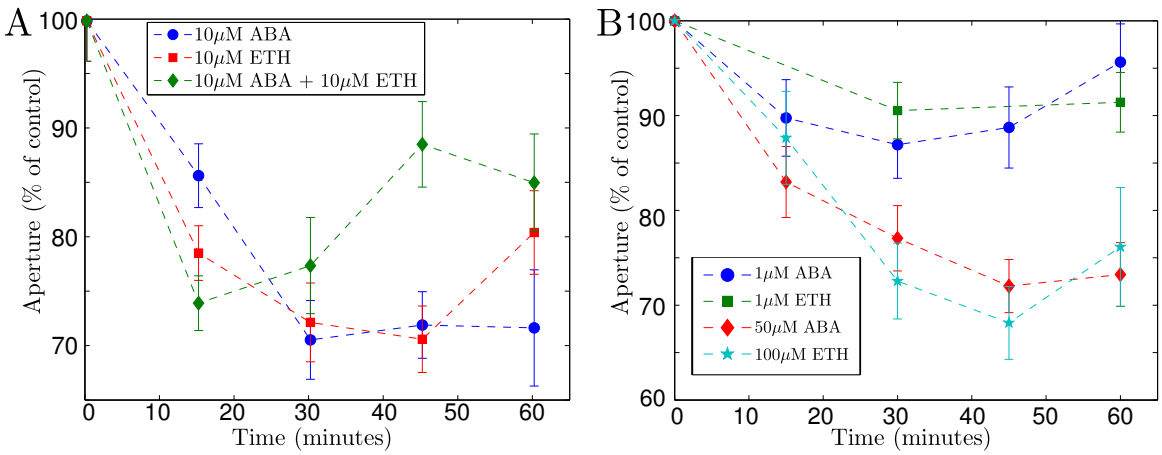

Figure 6.2: Stomatal aperture measurements as percentage of control. **A**: In response to 10 μM ABA (blue circles), 10 μM ethylene (red squares), and 10 μM ABA + 10 μM ethylene (green diamonds). **B**: In response to 1 μM ABA (blue circles), 1 μM ethylene (green squares), 50 μM ABA (red diamonds), and 100 μM ethylene (light-blue stars). Error bars show the standard error of the mean ($n = 30 \times 3$), control aperture is ≈ 1.9 μm.

to controls, in response to the treatments is shown in Fig. 6.1. ROS production is almost indistinguishable among treatments in the immediate five minutes after the stimuli. After 15 minutes, the cells stimulated only with ABA or ethylene still showed a 25% increase in their ROS concentration, whereas the increase in cells treated with a combined dose was only 5% above control. ABA and ethylene-induced ROS levels were maintained until 30 minutes after treatment, then a marked decrease was observed. After 60 minutes the ROS level in cells treated with ABA decreased but still remained higher than control, the cells treated with ethylene showed a reduction in ROS close to control levels, whereas the ROS level of the cells treated with the combined stimulus fell to 80% of control.

6.3 Aperture

The aperture responses of the guard cells treated with single and combined doses of 10μM ABA and 10μM ethylene are shown in Fig. 6.2A. As in other studies (Siegel *et al.*, 2009; Tanaka *et al.*, 2006) apertures are shown as percentage of control. The three treatments produced a decrease in stomatal aperture within 15 minutes of treatment. Interestingly, the cells with the combined ABA-ethylene dose showed an aperture reduction of 25%

compared to control, while cells treated with single stimuli of ABA and ethylene showed reductions of about 15% and 20% respectively. After 30 minutes the stomata with the combined treatment displayed less closure than the stomata under single treatments, a trend that was accentuated after 45 and 60 minutes (with a slight decrease at 60 min). Cells treated with ABA maintained their apertures close to 30% below control, whilst cells treated with ethylene had begun to re-open after 60 minutes but stayed below control.

Figure 6.2B shows additional measurements of stomatal closure made in cells stimulated with different individual doses of ABA and ethylene. From these data a similar behaviour to that in Fig. 6.2A can be observed, though with the additional information that the 1 μM doses are not sufficient to achieve full closure and the higher doses (50 μM and 100 μM) seem to have saturated. Note that the response of guard cells to 10 μM ethylene in Fig. 6.2A is consistent with the response to 100 μM ethylene in Fig. 6.2B causing its reopening 60 minutes after the stimulus.

6.4 Discussion

The time profiles of ROS production and stomatal aperture in response to single and combined 10 μM doses of ABA and ethylene shown in Fig. 6.1 and Fig. 6.2A are consistent with each other. It appears that a sustained elevation of ROS is needed to maintain stomata in a closed state. This is seen with individual ABA or ethylene stimulation, where up to 30 min there is a sustained level of ROS, coinciding with a decrease in aperture. Beyond this, for ABA there is still enough ROS up to 60 min to keep the stomata closed, but with ethylene, ROS levels appear to drop drastically, coinciding with a re-opening of stomata. In case of the combined stimulus, a substantial reduction in ROS coincides with a greater re-opening of the stomata.

The experimental observations in this chapter confirm the importance of ROS in ABA and ethylene-induced stomatal closure (as described in Chapter 2 and reported by Desikan *et al.* (2006); Kwak *et al.* (2003), and Pei *et al.* (2000)); they also provide new information about the mechanisms of guard cell signalling. From observing the ROS production in Fig. 6.1, it is immediately apparent that there is a mechanism only active under a simultaneous ABA and ethylene stimulus that removes hydrogen peroxide from the cells.

It is not clear what mechanisms are responsible for this behaviour, but as mentioned in Sec. 2.5, cross-talk has been known to occur in the ABA and ethylene signalling pathways, eg Wang *et al.* (2007). The interactions of the ABA and ethylene pathways seem to induce a specific antioxidant activity rapidly that is not present when the cell is presented with a single stimulus. A decrease in ROS levels between 30 and 60 minutes after treatment is observed in all treatments (Fig. 6.1). An antioxidant response to sustained high levels of ROS is unlikely to be the only mechanism present because cells with high levels of ROS (treated with ABA or ethylene individually) and lower ROS (treated with ABA and ethylene together) are affected similarly (Fig. 6.1). One explanation is that the individual stimuli, in parallel to stimulation of ROS production, also promote a delayed antioxidant response to allow ROS-mediated signalling to occur, removing ROS only after the signal has been transmitted. This delay is lacking in the combined stimulus situation, when another antioxidant mechanism is initiated early on. These results indicate that there may be at least two different antioxidant responses in guard cells to treatments of ABA and ethylene.

The observations reported here confirm the results of Tanaka *et al.* (2005) and expand them suggesting that complex interactions between the ABA and ethylene signalling pathways at the antioxidant level may be responsible for the reversal of stomatal closure following a combined stimulus. The long-term inhibition of stomatal opening seems unlikely to depend on sustained high ROS levels, due to the danger of oxidative stress. Results by Kim *et al.* (2010) and Pandey *et al.* (2007) indicate that long-term stomatal closure depends on different signalling components and gene-expression, mediated primarily by Ca^{2+}. It should be noted that, from the observations in Fig. 6.2A and B, ethylene does not seem to be able to keep stomata closed beyond the 60 minute mark. This reopening could be a distinctive feature of ethylene-induced stomatal closure that has so far eluded explanation. This hypothesis is further explored in Chapter 7 with the development of a differential equation model of stomatal closure that incorporates the antioxidant features that have been described here.

Chapter 7

Mathematical models of stomatal closure

The previous chapters discuss how unexpected outcomes may arise when multiple stimuli are applied to a large complex signalling network, even in well-studied systems such as guard cells, eg Desikan *et al.* (2006); Tanaka *et al.* (2005). In order to investigate the cause for the observed reversal of closure under combined ABA and ethylene stimuli, we develop a mathematical model for the signal transduction of these inputs in connection with stomatal closure. To construct the model we use the description of the signalling mechanisms involved in stomatal closure presented in Chapter 2, and the general mathematical modelling techniques presented in Chapter 3. The results and techniques of data fitting and signalling cascades from Chapters 4 and 5 are used to construct the model and fit its parameters. We begin by summarising and integrating the biology of the individual ABA and ethylene pathways to construct a single signalling network. Then, we make mathematical descriptions of the different signalling events of the network and, whenever possible, we seek expressions to simplify them. Finally, we fit the model to the experimental observations, analyse its results, and make predictions.

The ABA signalling network in guard cells has been studied computationally from an asynchronous dynamic boolean network perspective (Li *et al.*, 2006; Saadatpour *et al.*, 2010) and the ethylene pathway in root cells has been modelled using ODEs (Díaz and Álvarez-Buylla, 2006, 2009); however, to our knowledge there are no models of stomatal closure that incorporate both ethylene and ABA. We have chosen to construct our model using ODEs; motivated by the importance of the dynamical effects (as the experimental

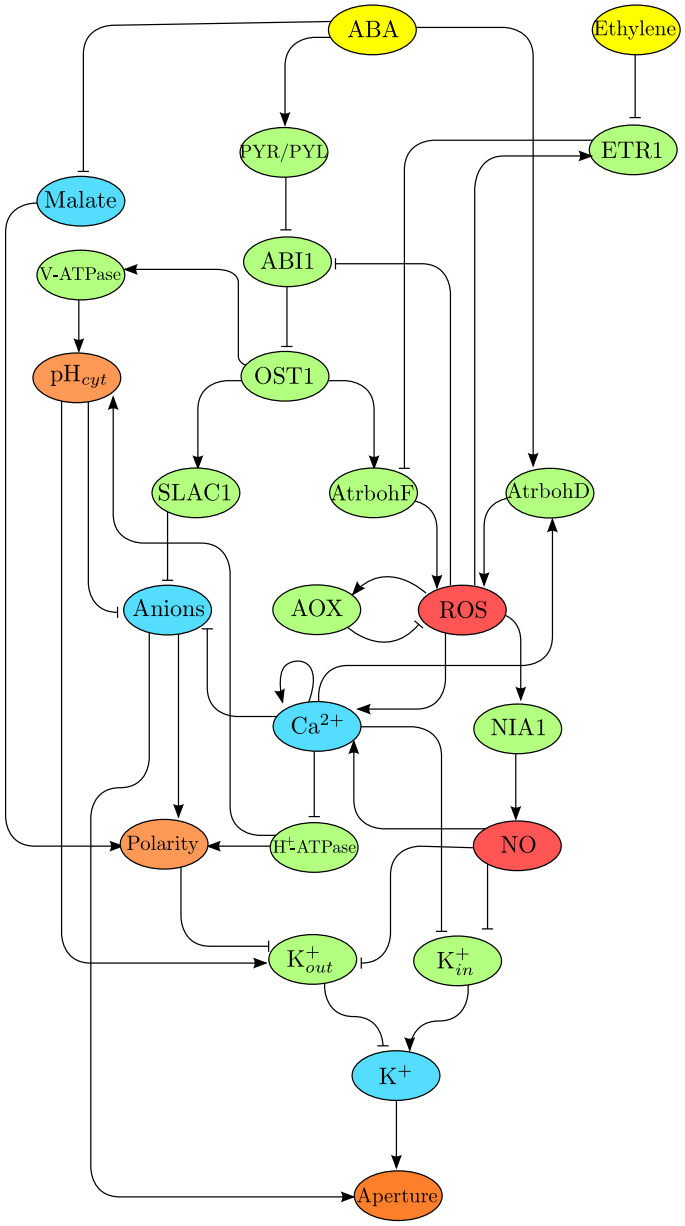

Figure 7.1: Integrated ABA and ethylene signalling network in guard cells. Interactions between the components of the network are shown by the edges joining them. Positive interactions such as activation or production are represented with edges ending in arrowheads $\rightarrow$. Negative interactions such as inactivation, repression, or scavenging are represented by edges ending in hammerheads $\dashv$. Yellow nodes are hormones, green nodes; proteins, blue nodes; ions, red nodes; reactive molecules, and orange nodes; physical properties of the cells. The interactions shown in this network are a summary of the signalling events described in Chapter 2.

results in Chapter 6 have shown) and the flexibility that such models afford to include the biochemical and biophysical interactions discussed in Chapter 2 which are summarised in Fig. 7.1 (see Sec. 7.1 below for a description). Conversely, we have aimed for a reduced model with a relatively small number of equations (variables) and parameters while still preserving dynamics and timescales. Such ODE models can be useful to test alternative biological hypotheses and can be amenable for more detailed analysis using bifurcation theory and comprehensive sensitivity analysis (August *et al.*, 2007).

7.1 Network representation of stomatal closure

The interactions of the components of the ABA and ethylene pathways described in Sec 2.3 and 2.4 are summarised in a network in Fig. 7.1. By constructing this network we inevitably make assumptions of what we consider are the most important components and interactions of the pathway, an inevitable consequence of the representation of current knowledge in a tractable way. In principle, each node in the network represents a component that could be a variable, either chemical, biochemical or physical, in a model. In the following sections we use the network representation of the pathway in Fig. 7.1 as a starting point to construct an ODE model of stomatal closure.

7.2 Construction of the model

7.2.1 Signal perception and ROS production

Sections 2.3.1, 2.3.2, and 2.4.2 describe the events in ABA and ethylene perception and production of ROS. These early events are represented in Fig. 7.2A by the subnetwork comprised of the nodes ABA, Ethylene, PYR/PYL, ABI1, OST1, AtrbohD/F, ROS, AOX and ETR1, whose concentrations we denote by $[ABA]$, $[ETH]$, $[PYR]$, $[ABI1]$, $[OST1]$, $[AtrbohF]$, $[AtrbohD]$, $[ROS]$, $[AOX]$, and $[ETR1]$. Given the timescales observed, we do not consider gene expression so the total amount of each enzyme in this subnetwork remains constant, thus we have the following conservation relations:

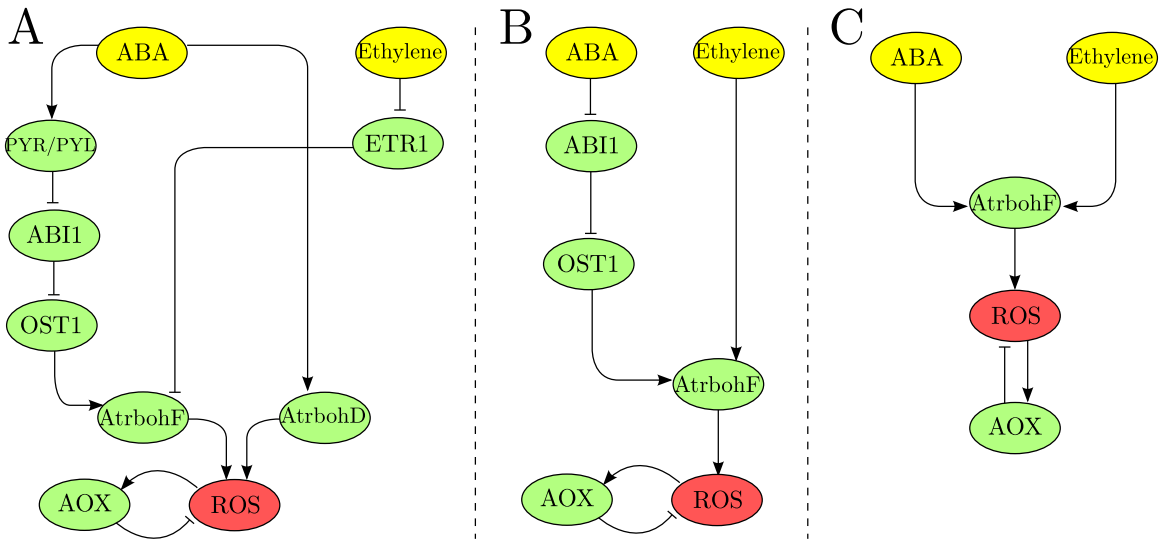

Figure 7.2: Simplification of signal perception and ROS production model. **A**: Network representation of the ROS production model (7.1)-(7.8). **B**: Simplification of the ROS production model given in equations (7.9)-(7.13). **C**: Further simplification of ROS production, given in equations (7.15)-(7.17).

$$
\begin{aligned}
[PYR_T] &= [PYR] + [PYR\text{-}ABA] + [PYR\text{-}ABI1], \\
[ABI1_T] &= [ABI1] + [ABI1\text{-}PYR], \\
[OST1_T] &= [OST1] + [OST1_P], \\
[AtrbohF_T] &= [AtrbohF] + [AtrbohF_P], \\
[AtrbohD_T] &= [AtrbohD] + [AtrbohD_P], \\
[ETR1_T] &= [ETR1] + [ETR1\text{-}ETH],
\end{aligned}
$$

where the name of the enzyme with a subscripted T denote total concentration of an enzyme (a non-negative constant in $\mathbb{R}$), those with a subscripted P; phosphorylated (or active, in general) enzymes, and variables joined with a dash are complexes. For example, the constant $[PYR_T]$ is the total amount of PYR/PYL in any form, the variable $[PYR]$ is the concentration of "available" PYR/PYL molecules, $[PYR\text{-}ABA]$ is the concentration of ABA-bound PYR/PYL molecules, and $[PYR\text{-}ABI1]$ is the concentration of ABI1-bound PYR/PYL. Based on the biology, we introduce the following model of what is known of

ABA and ethylene perception and ROS production:

$$\frac{\mathrm{d}[PYR\text{-}ABA]}{\mathrm{d}t} = k_1([PYR_T]\text{-}[PYR\text{-}ABA] + [PYR\text{-}ABI1])[ABA] - (k_{-1} + k_2([ABI1_T] - [ABI1\text{-}PYR]))[PYR\text{-}ABA], \tag{7.1}$$

$$\frac{\mathrm{d}[PYR\text{-}ABI1]}{\mathrm{d}t} = k_2[PYR\text{-}ABA]([ABI1_T] - [PYR\text{-}ABI1]) - k_3[PYR\text{-}ABI1], \tag{7.2}$$

$$\frac{\mathrm{d}[OST1_P]}{\mathrm{d}t} = k_4([OST1_T] - [OST1_P]) - (k_{-4} + k_5([ABI1_T] - [PYR\text{-}ABI1]]))[OST1_P], \tag{7.3}$$

$$\frac{\mathrm{d}[AtrbohF_P]}{\mathrm{d}t} = k_6([AtrbohF_T] - [AtrbohF_P])[OST1_P] - (k_7 + k_8([ETR1_T] - [ETR1\text{-}ETH])[AtrbohF_P], \tag{7.4}$$

$$\frac{\mathrm{d}[AtrbohD_P]}{\mathrm{d}t} = k_9([AtrbohD_T] - [AtrbohD_P])[ABA] - k_{10}[AtrbohD_P], \tag{7.5}$$

$$\frac{\mathrm{d}[ETR1\text{-}ETH]}{\mathrm{d}t} = k_{11}([ETR1_T] - [ETR1\text{-}ETH])[ETH] - k_{12}[ETR1\text{-}ETH], \tag{7.6}$$

$$\frac{\mathrm{d}[ROS]}{\mathrm{d}t} = k_{13} + k_{14}[AtrbohF_P] + k_{15}[AtrbohD_P] - (k_{16} + k_{17}[AOX])[ROS], \tag{7.7}$$

$$\frac{\mathrm{d}[AOX]}{\mathrm{d}t} = k_{18} + k_{19}[ROS] - k_{20}[AOX]. \tag{7.8}$$

Equations (7.1)-(7.3) describe the initial events of ABA perception, ABI1-phosphatase inhibition, and OST1 phosphorylation (Sec. 2.3.1). Equations (7.4) and (7.5) describe the activation of AtrbohD/F where OST1 phosphorylates AtrbohF, free ETR1 inactivates AtrbohF, and AtrbohD is activated by ABA (Sec. 2.3.2, and 2.4.2). Equation (7.6) shows ethylene binding and subsequent inactivation of ETR1 (Sec. 2.4.2). Equation (7.7) shows the production of ROS by AtrbohD/F and by other cellular processes (k_{13}), scavenging by antioxidants (summarised in the variable AOX), removal or decay of ROS ($k_{16}[ROS]$); equation (7.8) shows that AOX has endogenous production (k_{18}), production in response to ROS ($k_{19}[ROS]$), and decay or inactivation ($k_{20}[AOX]$).

Some things to note about the model above: the way in which the treatments are given (floated on a Petri dish, see Sec. 6.1) allows us to assume that the concentration of ABA and ethylene does not change, hence $[ABA]$ and $[ETH]$ are constants. The amount of ATP for the phosphorylation reactions and of NADPH for the production of ROS are also considered abundant enough within guard cells and they are implicitly included in the parameters of the model. We assume that when the complex PYR-ABA-ABI1 dissociates, it does so completely (ie into PYR, ABA, and ABI1), to make the equations simpler. The variable $[AOX]$ clusters together the group of antioxidants that are active during stomatal closure.

The model described in (7.1)-(7.8) has 8 equations and 20 kinetic parameters plus 6 parameters representing the total amount of the enzymes involved, which makes the

task of determining parameter values from the ROS measurements in Chapter 6 a rather difficult one. We attempt to reduce the number of equations and parameters as much as possible by making a series of assumptions, but avoiding oversimplification of the system. First, as noted by Kwak *et al.* (2003) and mentioned in Sec. 2.3.2 and Sec. 2.4.2, AtrbohD has a limited role in ABA-induced stomatal closure and none in ethylene-induced closure, so for our modelling we assume that ROS is exclusively produced by AtrbohF. This assumption is supported by the observations in Sec. 6.2 which show similar initial increases in ROS upon an ABA or ethylene stimulus (Fig. 6.1), for if there were two sources of ROS the pattern of initial increase would most likely differ. We make quasi-steady-state assumptions (QSSA) on the dynamics of the ABA and ethylene receptors (see Sec 3.2.2), so equations (7.1) and (7.6) are expected to reach equilibrium before the other variables have changed significantly. The expressions for the ligand-bound receptors then become:

$$[PYR\text{-}ABA] = \frac{([PYR_T] - [PYR\text{-}ABI1])[ABA]}{\frac{k_{-1}}{k_1} + \frac{k_2}{k_1}([ABI1_T] - [PYR\text{-}ABI1]) + [ABA]},$$
$$[ETR1\text{-}ETH] = \frac{[ETR1_T][ETH]}{\frac{k_{12}}{k_{11}} + [ETH]}.$$

The expression for $[ETR1\text{-}ETH]$ is a standard Michaelis-Menten term with a maximum rate of reaction given by the total amount of ETR1. The expression for $[PYR\text{-}ABA]$ is a Michaelis-Menten-type term with a dependency on the variable $[PYR\text{-}ABI1]$ (which holds PYR/PYL molecules). The system in (7.1)-(7.8) becomes

$$\frac{\mathrm{d}[PYR\text{-}ABI1]}{\mathrm{d}t} = \frac{k_2([ABI1_T] - [PYR\text{-}ABI1])([PYR_T] - [PYR\text{-}ABI1])[ABA]}{\frac{k_{-1}}{k_1} + \frac{k_2}{k_1}([ABI1_T] - [PYR\text{-}ABI1]) + [ABA]} - k_3[PYR\text{-}ABI1], \tag{7.9}$$
$$\frac{\mathrm{d}[OST1_P]}{\mathrm{d}t} = k_4(OST1 - [OST1_P]) - (k_{-4} + k_5([ABI1_T] - [PYR\text{-}ABI1]]))[OST1_P], \tag{7.10}$$
$$\frac{\mathrm{d}[AtrbohF_P]}{\mathrm{d}t} = k_6([AtrbohF_T] - [AtrbohF_P])[OST1_P] - \left(k_7 + k_8[ETR1_T]\left(1 - \frac{[ETH]}{\frac{k_{12}}{k_{11}} + [ETH]}\right)\right)[AtrbohF_P], \tag{7.11}$$
$$\frac{\mathrm{d}[ROS]}{\mathrm{d}t} = k_{13} + k_{14}[AtrbohF_P] - (k_{16} + k_{17}[AOX])[ROS], \tag{7.12}$$
$$\frac{\mathrm{d}[AOX]}{\mathrm{d}t} = k_{18} + k_{19}[ROS] - k_{20}[AOX]. \tag{7.13}$$

Figure 7.2B corresponds to this new model where some nodes and edges have been removed but the relationship between the signals and the output is still the same as in Fig. 7.2A. Though a simplification, this model is still large and we would like to find

a further simplification, taking into account what we have learnt so far. Note that in the network representation of equations (7.9)-(7.13) shown in Fig. 7.2B the path-length from ABA to AtrbohF is 3 while the path-length from ethylene to AtrbohF is 1; however, in Sec. 6.2 we observe a negligible difference between the ROS produced by ABA and the ROS produced by ethylene five minutes after treatment (Fig. 6.1). Note that the edge between ethylene and AtrbohF in Fig. 7.2B is positive because ETR1 inactivates the inactivator of AtrbohF. These observations suggest that the number of steps between ethylene perception and ROS production is about the same as with ABA signals (the immediate events after ethylene binding by ETR1 in guard cells are still unknown), or if the number of events is different then the timescale of the reactions is approximately the same in both cases. Furthermore, in both cases the maximum rate of ROS production is limited by the total amount of AtrbohF ($[AtrbohF_T]$), which means that the response to either signal has the same theoretical maximum rate. Given that data are unavailable for the receptors, OST1, and ABI1, and that ABA and ethylene have similar timescales for producing ROS we further simplify our model so that AtrbohF becomes active directly from the ABA and ethylene signals, as shown in Fig. 7.2C.

To determine the equations that represent the network in Fig. 7.2C, we must have a hypothesis of how ABA and ethylene signals activate AtrbohF. One possibility is that the signals activate AtrbohF through the same pathway, ie there is a bottleneck for both signals upstream of AtrbohF. In this case the signals are essentially interchangeable. This assumption would imply that, for example, a 2 μM dose of ABA and a combined 1 μM ABA plus 1 μM ethylene dose are the same:

$$\frac{\mathrm{d}[ROS]}{\mathrm{d}t} \approx k_{13} + \frac{[AtrbohF_T]([ABA + ETH])}{\kappa + ([ABA + ETH])} - (k_{16} + k_{17}[AOX])[ROS].$$

The identity of the signalling "bottleneck" remains unknown. Another hypothesis that has a similar result but that does not require a common node in the pathways of the signals would be to assume that the signals converge for the first time at AtrbohF, and activate it independently of each other:

$$\frac{\mathrm{d}[ROS]}{\mathrm{d}t} \approx k_{13} + \frac{\alpha_1\kappa_2[ABA] + \alpha_2\kappa_1[ETH]}{\kappa_1\kappa_2 + \kappa_2[ABA] + \kappa_1[ETH]} - (k_{16} + k_{17}[AOX])[ROS], \quad (7.14)$$

where α_1 is a product of $[AtrbohF_T]$ and other rate-limiting parameters in the ABA pathways, and κ_1 is the ABA-specific Michaelis constant, and likewise for ethylene α_2 is the rate-limiting parameter and κ_2 the Michaelis constant. We use the approximation ($\approx$) sign to emphasise that this is not a rigorous derivation of the ROS-activation dynamics but a deduction guided by our current knowledge of the system and assumptions deemed reasonable. In Appendix A.2.2 we derive the compound Michaelis-Menten term in equation (7.14) using the QSSA. Though both hypotheses of ROS production can produce a similar response the latter one is better suited for modelling stomatal closure because it does not require the assumption of additional interactions, includes the former as a special case (when $\kappa_1 = \kappa_2$ and $\alpha_1 = \alpha_2 = [AtrbohF_T]$), and its derivation is clear.

Now we turn our attention to the antioxidant pool AOX; it is clear that a homogeneous antioxidant pool responsive only to the concentration of ROS as described by equation (7.13) is incompatible with our observations in Chapter 6. Following the experimental indications in Sec. 6.2, we consider the possibility of two different antioxidant mechanisms described by the variables $[AOX_1]$ and $[AOX_2]$ which lie at the end of linear activation cascades driven by $[ABA]$ and $[ETH]$ (Fig. 7.3). As discussed when the experimental results were presented, there is evidence to suggest that two distinct antioxidant mechanisms might be at work during stomatal closure. The first of these mechanisms (AOX_1) can be activated by either an ABA *or* an ethylene signal and includes endogenous antioxidant production (to maintain unstimulated equilibrium levels), whereas the second antioxidant (AOX_2) is active only when ABA *and* ethylene signals are present simultaneously. Logical *and* and *or* gates in biochemical systems can be the result of particular post-translational modification of enzymes, eg multiple phosphorylation, trimerisation, etc (Mayo *et al.*, 2006).

We place the response of the antioxidant to the signals at the end of cascades of abstract variables to emulate the delay observed in the removal of ROS. Each cascade has a constant input (depending on the ABA and ethylene signals, see below) and has a solution proportional to the lower-incomplete Gamma function $\mathrm{P}(n_i, h_i t)$ (see Sec. 5.4.3 and Appendix C.1.2) where n_i represents the length of the cascade and h_i is related to the deactivation rates of the cascade, $i = 1, 2$.

The input of the cascade culminating in AOX_1 must follow a boolean *or* logic and

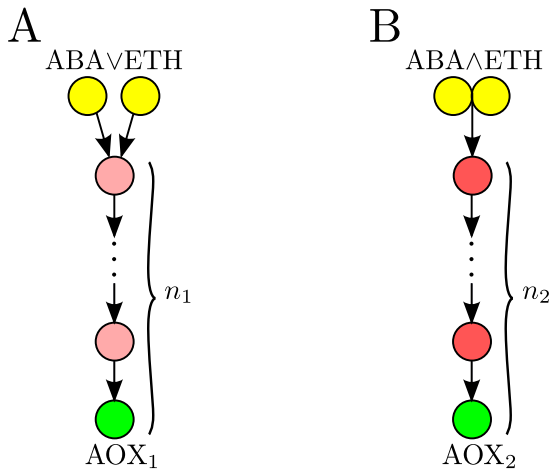

Figure 7.3: Antioxidant production at the end of linear activation cascades. **A**: Cascade of length n_1 whose input has a logic *or* function, the cascade responds to ABA or ethylene treatment (equation (7.16)). **B**: Cascade of length n_2 whose input has a logic *and* function, the cascade becomes active only when ABA and ethylene are present simultaneously (equation (7.17)).

saturate, therefore we use the compound Michaelis-Menten form we used to describe ROS-production previously as the input for the ABA and ethylene signals, so the equation for $[AOX_1]$ is

$$\frac{\mathrm{d}[AOX_1]}{\mathrm{d}t} = k_{18} + \frac{\alpha_3\kappa_4[ABA] + \alpha_4\kappa_3[ETH]}{\kappa_3\kappa_4 + \kappa_4[ABA] + \kappa_3[ETH]}\mathrm{P}(n_1, h_1 t) - k_{20}[AOX_1].$$

The cascade leading to the activation of AOX$_2$ must operate as a logic *and* gate, ie becoming active only if $[ABA] > 0$ and $[ETH] > 0$. We consider that the input of the cascade is downstream of the ABA and ethylene receptors and the response of the cascade must also exhibit saturation, so we model the input as the product of two Michaelis-Menten forms:

$$\frac{\alpha_5[ABA]}{\kappa_5 + [ABA]}\frac{\alpha_6[ETH]}{\kappa_6 + [ETH]},$$

which enforce a logical *and* operation of the signals. The expression for AOX$_2$ is

$$[AOX_2](t) = \frac{\alpha_5\alpha_6[ABA][ETH]}{(\kappa_5 + [ABA])(\kappa_6 + [ETH])}\mathrm{P}(n_2, h_2 t).$$

Now we have a complete reduced model of a ROS production-module in guard cells:

$$\frac{\mathrm{d}[ROS]}{\mathrm{d}t} = k_{13} + \frac{\alpha_1\kappa_2[ABA] + \alpha_2\kappa_1[ETH]}{\kappa_1\kappa_2 + \kappa_2[ABA] + \kappa_1[ETH]} - (k_{16} + k_{17}[AOX_1] + k_{21}[AOX_2])[ROS], \tag{7.15}$$

$$\frac{\mathrm{d}[AOX_1]}{\mathrm{d}t} = k_{18} + \frac{\alpha_3\kappa_4[ABA] + \alpha_4\kappa_3[ETH]}{\kappa_3\kappa_4 + \kappa_4[ABA] + \kappa_3[ETH]}\mathrm{P}(n_1, h_1 t) - k_{20}[AOX_1], \tag{7.16}$$

$$[AOX_2](t) = \frac{\alpha_5\alpha_6[ABA][ETH]}{(\kappa_5 + [ABA])(\kappa_6 + [ETH])}\mathrm{P}(n_2, h_2 t), \tag{7.17}$$

with initial conditions $[ROS](0) = [ROS]_0$, $[AOX_1](0) = [AOX_1]_0$, and $[AOX_2](0) = 0$ to be determined at a later time.

7.2.2 NO production

A description of how ROS induces NO production in guard cells treated with ABA via the enzyme NIA1 is given in Sec. 2.3.3 (Fig. 7.4A); and in Sec. 2.4.3 preliminary indications that NO is also produced in guard cells treated with ethylene were discussed.

An initial ODE describing endogenous and enzymatic NO production in guard cells is

$$\frac{\mathrm{d}[NO]}{\mathrm{d}t} = \alpha_{30} + \frac{\alpha_{31}[ROS]}{k_{31} + [ROS]} - \beta_{30}[NO],$$

where α_{30} is a constant rate of NO production by other processes, the Michaelis-Menten term is ROS-induced NO production via NR1, and the last term is NO decay and removal.

Note that in unstimulated guard cells $[ROS] > 0$ and as measurements of NO to distinguish between ROS and non-ROS induced production are not available, we gather the ROS-dependent and ROS-independent NO production in a single term. We include a second term that describes further enzymatic NO production from ethylene (Fig. 7.4B), which could be either from NIA1 or another, yet unidentified source. This term is needed by our fits to the data, as preliminary models without it were unable to reproduce experimental observations. The new expression for NO production becomes

$$\frac{\mathrm{d}[NO]}{\mathrm{d}t} = \frac{\alpha_{31}[ROS]}{k_{31} + [ROS]} + \frac{\alpha_{32}[ETH]}{k_{32} + [ETH]} - \beta_{30}[NO], \tag{7.18}$$

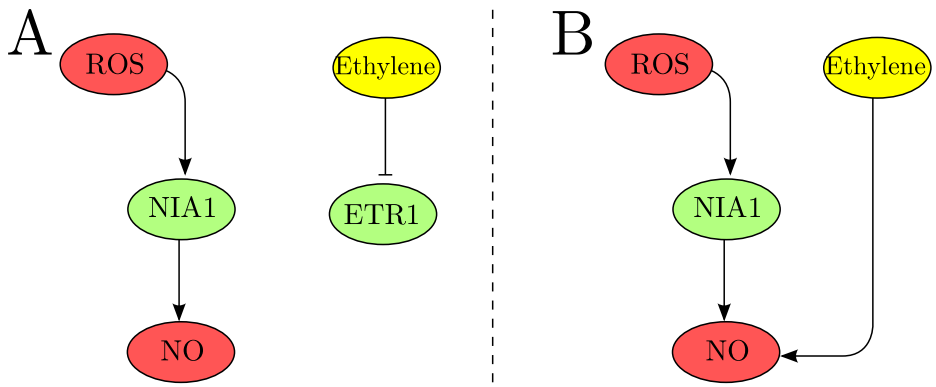

Figure 7.4: NO production models in guard cells. **A**: NO-production network described in Sec. 2.3.3 and Sec. 2.4.3. NO is produced in a ROS-dependent way by NIA1. **B**: Model of NO-production given in equation (7.18). The key assumption of this model is the existence of a ROS-independent pathway of NO production in response to ethylene.

with initial condition $[NO](0) = [NO]_0$, which is to be determined later.

7.2.3 Ca^{2+} increase, cytosolic alkalinisation, and ion efflux

As discussed in Sec. 2.3.5, the dynamics of Ca^{2+}-release and action in guard cells are complex and not fully understood. Though the importance of Ca^{2+} in guard cell signalling (and cell viability in general) is beyond doubt, in this work we do not include an equation describing its behaviour for three reasons:

i. Reports of Ca^{2+}-behaviour after ABA treatments in the literature describe both oscillations and rises in cytosolic levels, and experimental data-sets encompassing both single and combined ABA and ethylene treatments do not exist.

ii. As mentioned by Kim *et al.* (2010), the way in which a cytosolic Ca^{2+} rise (or oscillations) transmit signals during stomatal closure is not yet clear. Current hypotheses state that ABA "primes" Ca^{2+} receptors, making the rise in cytosolic levels helpful but not essential for successful closure.

iii. Experimental evidence presented in Chapter 6 suggests that ABA-ethylene cross-talk occurs at the ROS-level, upstream of Ca^{2+} in the guard cell signal transduction network in Fig. 7.1, so we direct most of our efforts to understanding signal

transduction at this level.

In Sec. 2.3.4 we described the process of cytosolic alkalinisation whereby the pH in the cytosol of guard cells increases from 7.0 to 7.5, and the pH in the tonoplast decreases from 5.5 to 5.0 after treatment with ABA. The concentration of H^+ determines pH (see Appendix D.1 for a brief introduction to the Henderson-Hasselbach equation): $pH = -\log([H^+])$. During stomatal closure the membrane H^+–ATPases are inactivated, which means that changes in pH are the result of the transport of protons from the cytosol into the tonoplast by the vacuolar proton pumps (V–ATPases), activated by OST1 (Fig. 7.5A, see Appendix D.2 for a brief introduction to proton transport by ATPases).

The model of cytosolic alkalinisation and ion-efflux shown in Fig. 7.5B has one equation for the potassium ion concentration $[K^+]$, and one equation for the outwards potassium channels $[K^+_{out}]$:

$$\frac{\mathrm{d}[K^+_{out}]}{\mathrm{d}t} = \alpha_{40} + \frac{\alpha_{41}[ABA]}{k_{11} + [ABA]} + \alpha_{42}[NO] - \beta_{40}[K^+_{out}], \tag{7.19}$$

$$\frac{\mathrm{d}[K^+]}{\mathrm{d}t} = \frac{\alpha_{51}}{k_{51} + [NO]} - \beta_{50}[K^+_{out}][K^+]. \tag{7.20}$$

Equation (7.19) shows the change in $[K^+_{out}]$, the active outwards K^+ channels. The first (α_{40}) and last ($\beta_{40}[K^+_{\mathrm{out}}]$) terms represent the constant flux of channels between the active and inactive states, respectively. The second term represents the extra number of channels made available by the increase in cytosolic pH following an ABA stimulus, mediated by OST1. The third term ($\alpha_{42}[NO]$) is the increase in K^+_{out} activity as a result of membrane depolarisation, possibly via NO-induced Ca^{2+} release (ie via the path NO $\rightarrow$ Ca^{2+} $\dashv$ H^+–ATPase $\rightarrow$ Polarity $\dashv$ K^+_{out} in Fig. 7.5A). Here we assume that NO does not target K^+_{out}, as ion efflux is required for stomatal closure (we note that although NO has been shown to block K^+_{out} in *Vicia faba* guard cells (Sokolovski and Blatt, 2004), the authors of the study are unsure whether NO action is specifically targeted to K^+_{out}.) Equation (7.20) shows the change in $[K^+]$. The first term represents the increase of ions that enter through the inwards-rectifying channels (K^+_{in}), which are inactivated by NO. The second term is the ion efflux through the outwards channels that is proportional to the active channels $[K^+_{out}]$ and the ion concentration itself. We include an equation for K^+_{out} but not for K^+_{in} because the

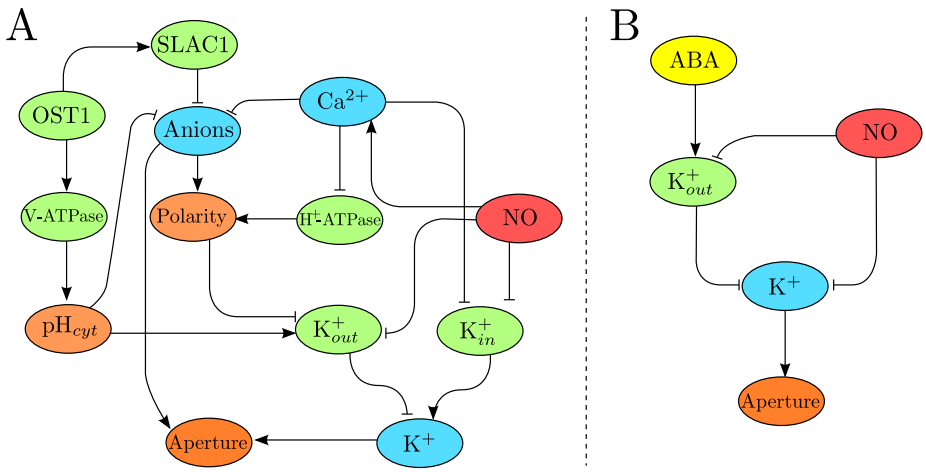

Figure 7.5: Events in stomatal closure downstream of ROS. **A**: Cytosolic alkalinisation following treatment with ABA, Ca^{2+}–increase, membrane depolarisation, and ion efflux. **B**: Model of late events in stomatal closure presented in equations (7.19) and (7.20).

alkalinisation of the cytosol has the effect of increasing the number of available channels to extrude ions, whereas the inactivation of K^+_{in} is only represented by a term in the equation for K^+.

7.2.4 Loss of turgor

The relationship of this model to stomatal aperture is via the last variable $[K^+]$. Cell volume (and hence stomatal aperture) is determined by the ion and solute concentration in the cell relative to the external concentration (Pandey *et al.*, 2007; Tanaka *et al.*, 2006). Therefore, we take ions and solutes (K^+ in particular) as a simple proxy for aperture:

$$[AP] \propto [K^+]. \tag{7.21}$$

7.3 A model of signal transduction for stomatal closure

We use equations (7.15)-(7.20) to construct a a model of ABA and ethylene-induced stomatal closure which we represented graphically in Fig. 7.6. The model describes the dynamics of six variables in terms of the external input functions $[ABA]$ and $[ETH]$, which denote the doses of the treatments. We normalise the variables in the model by their non-stimulated equilibrium levels (ie the initial conditions) so they represent

percentage of control:

$$\widehat{[ROS]}(t) = \frac{100[ROS](t)}{[ROS]_0},$$
$$\widehat{[AOX_1]}(t) = \frac{100[AOX_1](t)}{[AOX]_0},$$
$$\widehat{[NO]}(t) = \frac{100[NO](t)}{[NO]_0},$$
$$\widehat{[K^+_{out}]}(t) = \frac{100[K^+_{out}](t)}{[K^+_{out}]_0},$$
$$\widehat{[K^+]}(t) = \frac{100[K^+](t)}{[K^+]_0},$$
$$\widehat{[AP]}(t) = \frac{100[AP](t)}{[AP]_0}.$$

With the normalised variables we can transform equation (7.21) to $\widehat{[K^+]}(t) = \widehat{[AP]}(t)$, so we take the normalised potassium ion concentration as equivalent to the normalised aperture. Dropping the hat notation, the equations of the model become (note that the parameters have been renamed)

$$\frac{\mathrm{d}[ROS]}{\mathrm{d}t} = \alpha_{10} + \frac{\alpha_{11}k_{12}[ABA] + \alpha_{12}k_{11}[ETH]}{k_{11}k_{12} + k_{12}[ABA] + k_{11}[ETH]} - (\beta_{11}[AOX_1] + \beta_{12}[AOX_2])\,[ROS], \quad (7.22)$$

$$\frac{\mathrm{d}[AOX_1]}{\mathrm{d}t} = \alpha_{20} + \frac{\alpha_{21}k_{22}[ABA] + \alpha_{22}k_{21}[ETH]}{k_{21}k_{22} + k_{22}[ABA] + k_{21}[ETH]}\,\mathrm{P}(n_1, \alpha_{23}t) - \beta_{20}[AOX_1], \quad (7.23)$$

$$[AOX_2](t) = \frac{[ABA][ETH]}{(k_{11} + [ABA])(k_{12} + [ETH])}\mathrm{P}(n_2, \beta_{13}t). \quad (7.24)$$

$$\frac{\mathrm{d}[NO]}{\mathrm{d}t} = \frac{\alpha_{31}[ROS]}{k_{31} + [ROS]} + \frac{\alpha_{32}[ETH]}{k_{12} + [ETH]} - \beta_{30}[NO], \quad (7.25)$$

$$\frac{\mathrm{d}[K^+_{\mathrm{out}}]}{\mathrm{d}t} = \alpha_{40} + \frac{\alpha_{41}[ABA]}{k_{11} + [ABA]} + \alpha_{42}[NO] - \beta_{40}[K^+_{\mathrm{out}}], \quad (7.26)$$

$$\frac{\mathrm{d}[K^+]}{\mathrm{d}t} = \frac{\alpha_{51}}{k_{51} + [NO]} - \beta_{50}[K^+_{\mathrm{out}}][K^+]. \quad (7.27)$$

The model describes the dynamics of six variables in terms of the external input functions $[ABA]$ and $[ETH]$, which denote the doses of ABA and ethylene respectively. Note that the Michaelis-Menten term involving ethylene in equation (7.25) has the same Michaelis constant k_{12} associated with $[ETH]$ in equation (7.22). The term with ABA in equation (7.26) has also the same constant k_{11} as the one associated with $[ABA]$ in the ROS equation.

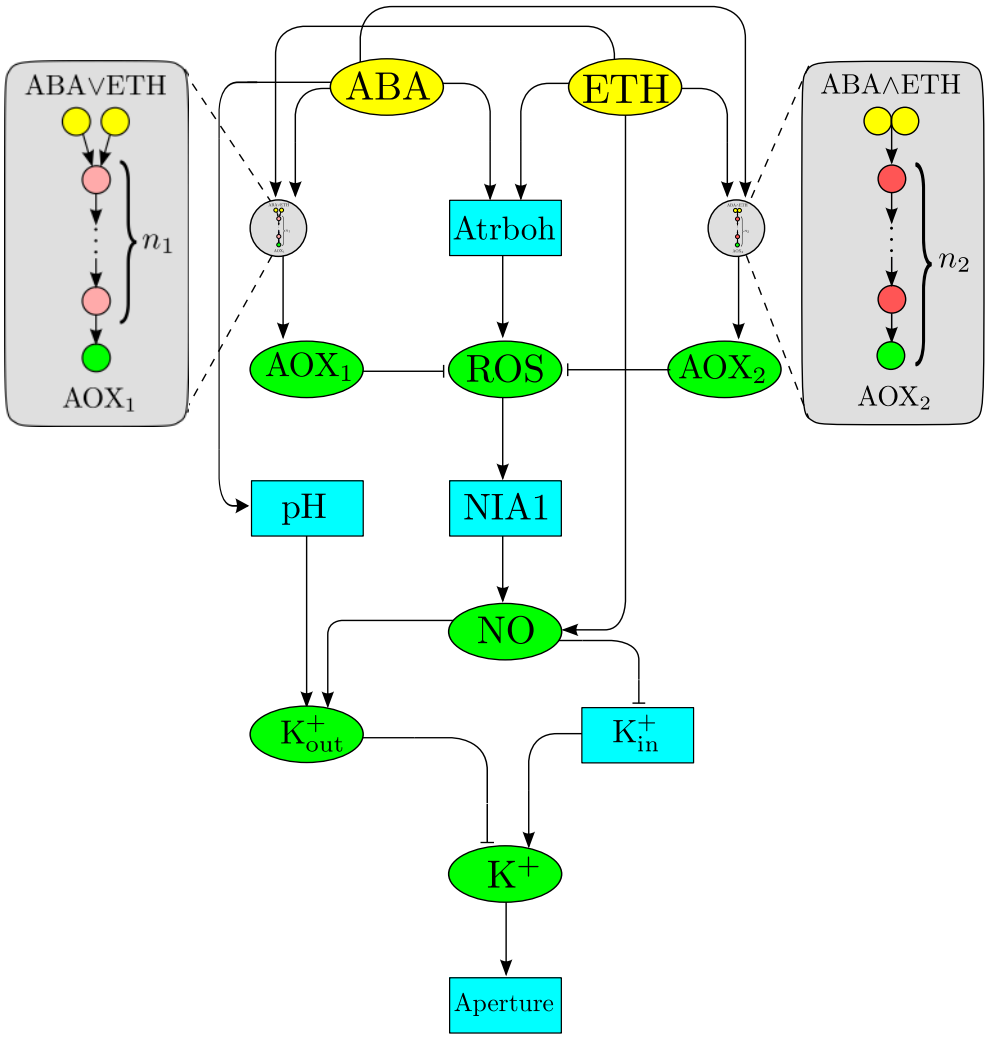

Figure 7.6: Model of stomatal closure under ABA and ethylene stimuli. All nodes in this picture are either explicitly or implicitly included in equations (7.7)-(7.20). The input nodes are shown as yellow ellipses, variable nodes are shown in green ellipses, and nodes that are blue rectangles are implicitly represented in the equations. The two circular grey nodes represent parts of the signal transduction network whose components are not yet known, and we include them as linear activation cascades (see text). The cascade AOX_1 (on left) is activated by performing a logical *or* operation on the ABA and ethylene signals, the presence of either suffices to elicit a response. The cascade upstream of AOX_2 (on right) becomes activated by performing a logical *and* operation on the ABA and ethylene signals; the presence of both of them is required to elicit a response.

One of the distinctive aspects of this model is the approach towards the simplification of uncharacterised activation cascades, which can control the strength and timing of the antioxidant response (Heinrich *et al.*, 2002). In particular, it is key that the modelling of the cascades in this model incorporates a representation of the implicit delay present in the antioxidant response, which occurs in parallel to the production of ROS (see the cascades on both sides of Fig. 7.6). In order to accomplish this parsimoniously within the setting of ODEs, equations (7.23) and (7.24) incorporate cascading terms which introduce a delay through the effect of parallel processes. The introduction of such terms is commonplace as an alternative to more complex delay equations (Bar-Or *et al.*, 2000; Höfer *et al.*, 2002). In our case, we have used a simple model of a linear activation cascade with identical deactivation rates (Heinrich *et al.*, 2002), which has been shown to provide optimal signal amplification (Chaves *et al.*, 2004). Each cascading module has an explicit analytical solution in terms of the normalised incomplete gamma function and introduces only three parameters to the model, as discussed in Chapter 5.

7.3.1 Numerical results from the model

The resulting dynamical model of signal transduction (7.22)–(7.27) has six variables and consists of five coupled nonlinear ODEs (because one of the variables can be solved explicitly as a function of time) with 28 parameters. We now fit the model to experimental data and we use it to study the temporal response of the system to external inputs of $[ABA]$, $[ETH]$ and combined $[ABA] + [ETH]$ stimuli.

Fitting the model to experimental observations

Our simulations start from a 'control' initial condition:

$$[ROS](0) = [AOX_1](0) = [NO](0) = [K^+_{out}](0) = [K^+](0) = 100,$$
$$\text{and } [AOX_2](0) = 0. \tag{7.28}$$

Furthermore, we define relationships between some of the parameters such that when there is no treatment (ie $[ABA] = [ETH] = 0$), the system remains in equilibrium at the control

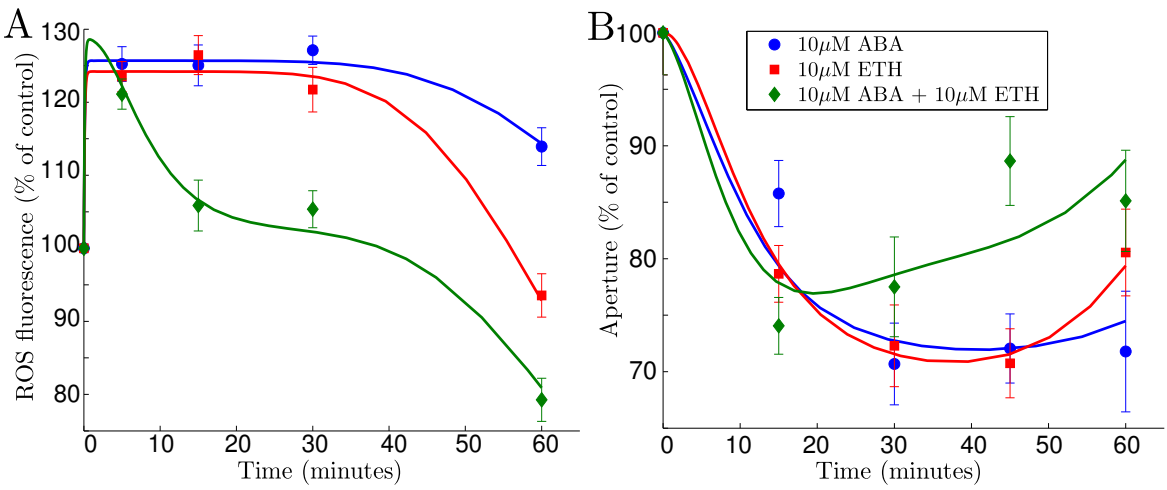

Figure 7.7: Experimental data and model results for 10 μM doses. **A**: ROS fluorescence time-course. **B**: Stomatal aperture time-course. Plots show responses after treatment with 10 μM ABA (blue), 10 μM ethylene (red), and 10 μM ABA and 10 μM ethylene (green). Markers show experimental measurements presented in Chapter 6, bold lines are the solutions to equations (7.22)–(7.27) with the parameter values from Table D.1.

initial condition.

We use the experimental time-course measurements of ROS and aperture presented in Chapter 6 to fit the model parameters. For a relatively large model with many parameters, the amount of data we have is not very much, which will make parameter fitting a challenging task. We use the *squeeze-and-breathe* optimisation procedure introduced in Chapter 4 to find the values of the parameters. The squeeze-and-breathe procedure is suited for optimisation problems such as this, given its ability to navigate the complex surfaces of the objective function (the difference between the model and the data as a function of the parameters) and find its minimum. Appendix D.3 gives the details of the implementation. The parameters of the model found using the squeeze-and-breathe algorithm are shown in Table D.1.

The fits to the data in Fig. 7.7 are shown to match the set of observations after single and combined 10 μM treatments, specifically the response to the combined stimuli. Figure 7.8 shows solutions of the model along with the aperture observations under assorted treatments; these show that the model is able to reproduce observations that range from low-dose treatments to higher doses.

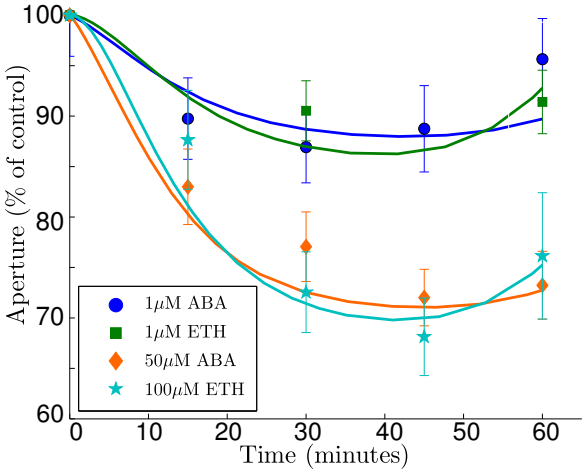

Figure 7.8: Experimental aperture data and model results. Stomatal aperture after treatment with 1 μM ABA (blue circles), 1 μM ethylene (green squares), 50 μM ABA (orange diamonds), and 100 μM ethylene (light-blue stars). Bold lines show the solutions to equations (7.22)–(7.27) with the parameters in Table D.1.

7.3.2 Dynamical response of the model to stimuli

The dynamical behaviour of the model of signal transduction is explored in Figures 7.9 and 7.10. We remark again that the model is constructed to represent only the transient dynamics following different external inputs and that we have not considered further downstream mechanisms that would dominate the dynamics at longer timescales. Stomatal closure in response to either ABA or ethylene is a relatively fast process that takes place in time scales shorter than typical genetic regulation. Maintaining stomatal closure and inhibiting stomatal opening are separate processes (Pandey *et al.*, 2009), that require other regulatory interactions and expression of certain genes (Cho *et al.*, 2009a; Kim *et al.*, 2010) which we do not consider here. Therefore, we are only concerned in this work with short-term, transient behaviour of stomata and we do not study the stationary dynamics of the model. Once the model reaches steady state, other processes will be active in guard cells affecting the model behaviour.

The heat maps in Fig. 7.9 show snapshots of the time-course of stomatal aperture following treatments of different dose combinations of ABA and ethylene, represented on the $([ABA], [ETH])$ plane. The simulations reproduce the observation that doses of combined treatment (up to 20 μM) result in diminished closure (that could lead to

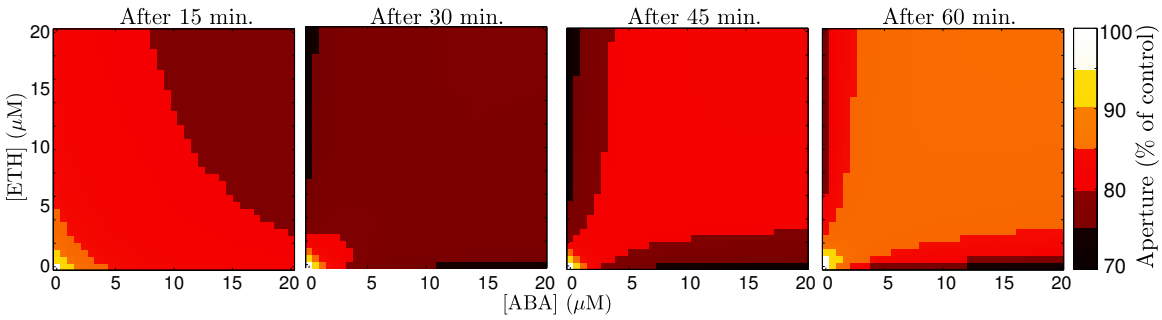

Figure 7.9: Predicted aperture response map of combinations of ABA and ethylene doses 15, 30, 45, and 60 minutes after treatment. The coordinates of each point in represent the dose combination and the colour denotes the response. Dark regions indicate less aperture (ie more closure in response to treatment) and lighter regions show more aperture. Response apertures shown as percent of control.

enhanced aperture if enough ROS is depleted) over the time course, as compared to the increased closure induced by doses of single treatments. The variables respond with different intensity to the treatments; in particular the model predicts that ethylene has a stronger $[AOX_1]$ response than ABA (see Fig. 7.11A whose consequence can be observed in the faster decay in ROS fluorescence from Fig. 7.7A). The antioxidant $[AOX_2]$ is only active only during treatment with both hormones, as we have previously discussed.

Figure 7.10 shows that the responses to the two hormones (after 60 minutes) are asymmetrical, with ABA inducing more closure than ethylene. This asymmetry of the response is consistent with reports in the literature that stomata respond more strongly to ABA than to ethylene (Tanaka *et al.*, 2005), though it appears to diminish with increased doses.

A prediction of the model is that relatively low doses of the combined stimulus "backfire" 45 minutes after treatment, resulting in the arrest of the closure process. A slight reopening can be observed in the response to combined treatments in Fig. 7.10; such re-opening is a direct consequence of the excess antioxidant activity (ie $[AOX_2]$) that results from the interaction of the ABA and ethylene pathways (Fig. 7.11). Further experimental data are required to validate this prediction and to check if other mechanisms become significant at longer times.

The modelling also suggests that ethylene should have more than one pathway to produce NO. Alternative models with ethylene-induced NO exclusively produced via ROS

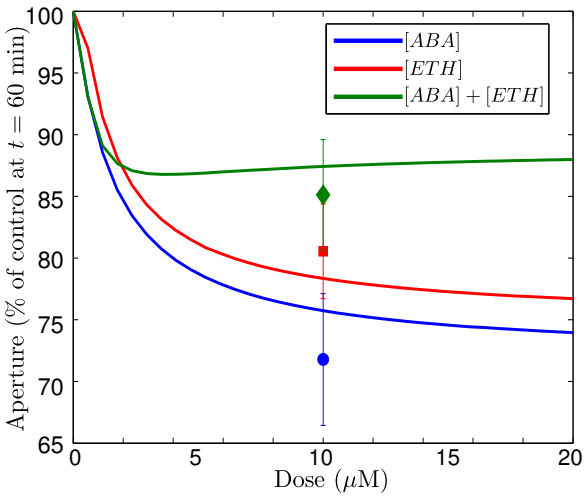

Figure 7.10: Predicted aperture response after 60 minutes to doses of ABA (blue line), ethylene (red line), and a combination of equal doses of ABA and ethylene (green line). The responses to single treatments correspond to the x and y axes of Fig. 7.9. The green line is the response to the combined stimulus of equal doses of ABA and ethylene, which corresponds to the values along the diagonal in the heat map. The data points are the experimental measurements at $t = 60$ minutes of the aperture also shown in Fig. 7.7.

(ie by making $\alpha_{32} = 0$) were not able to reproduce the response dynamics of guard cells that we report here. In parallel experimental studies in R. Desikan's lab, new signalling pathways downstream of ethylene that appear to be ROS-independent are being uncovered. Similarly with the active K^+_{out} channels: ABA-driven alkalinisation alone is not enough to create the outwards flux of ions needed to achieve stomatal closure and an NO term ($\alpha_{42}[NO]$) must be added to equation (7.26) to attain the necessary ion flux. The relationship between NO and K^+_{out} is unlikely to be direct, although NO can block K^+_{out} by nitrosylation (Sokolovski and Blatt, 2004). Enhancement of K^+_{out} activity by NO is more likely driven by membrane depolarisation in response to Ca^{2+} release or cytosolic alkalinisation, as depicted in Fig. 7.1.

Using our analysis of DDEs from Sec. 5.4.3 we can calculate the characteristic timescales τ_1 and τ_2, that lead to the activation of AOX_1 and AOX_2 respectively:

$$\tau_1 = \frac{n_1}{\alpha_{23}} - 1 \approx 111 \text{ min}, \qquad \tau_2 = \frac{n_2}{\beta_{13}} - 1 \approx 12 \text{ min}.$$

These timescales (visible in the time-courses in Fig. 7.11) give information about the

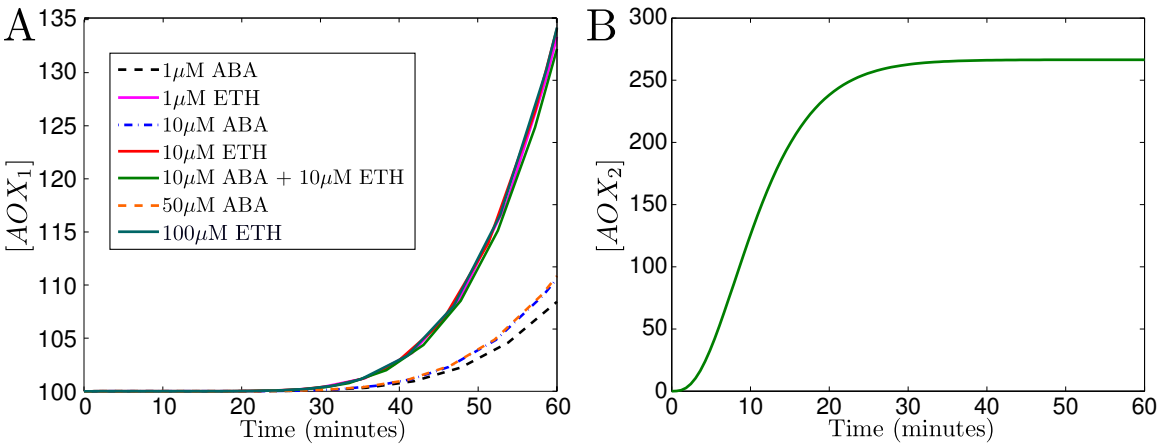

Figure 7.11: Antioxidant time-courses after treatment. **A**: Time-course of AOX_1 given by equation (7.23), responses to ABA treatments (dashed-lines) are lower than the responses to treatments which include ethylene. **B**: AOX_2 time-course given by equation (7.24), only active when ABA and ethylene treatments are present at the same time.

nature of the mechanisms behind the observed antioxidant effects. The value of the AOX_1 timescale τ_1 suggests there is enough time for antioxidant genes to be expressed. In contrast τ_2 is smaller, meaning that AOX_2 has a much faster timescale and its action is unlikely to depend on gene expression.

7.4 Discussion

In this chapter, we have investigated theoretically why full stomatal closure fails to occur when guard cells are presented with a combined ABA-ethylene stimulus, an observation first reported by Tanaka *et al.* (2005), and reproduced in Chapter 6. As shown in Fig. 7.1, the pathways of both hormones overlap strongly, with ROS playing a significant role. The development of earlier models that culminated in the model presented here, has required information about ROS levels past the 5 minute mark, after treatment, hence the experimental observations of ROS production have been extended to 60 minutes. The measurements of ROS and stomatal aperture in guard cells stimulated with ABA, ethylene, and ABA plus ethylene show that when both hormones are present, ROS are removed swiftly after an initial burst of production and the closure process reverses. This is the first report of a rapid change and shift in pattern of ROS production in guard cells depending on

the type and number of input stimuli.

Based on the experimental data, the model presented in this chapter posits the existence of two separate antioxidant mechanisms active in guard cells. One such mechanism is the generic antioxidant mechanism (AOX_1) operational with any single ROS-generating stimulus (ABA *or* ethylene), which allows ROS to signal downstream components of the pathway and then removes ROS to control oxidative stress. The second antioxidant (AOX_2) response is only active when both hormones, ABA *and* ethylene, are present simultaneously and does not allow the ROS signal to persist long enough to maintain closure, causing the reversal of the closure process.

The antioxidant pool in guard cells is diverse, including eg ascorbate, catalase, and glutathione (Chen and Gallie, 2004; Pham and Desikan, 2009). Furthermore, NO has been shown to exhibit antioxidant activity by reacting with superoxide (Neill *et al.*, 2008), to enhance dessication tolerance (Bai *et al.*, 2011), and to nitrosylate NADPH-oxidase (Yun *et al.*, 2011). Although the interactions between ABA, ethylene, signalling molecules and antioxidants are highly complex, the model developed here presents a first hypothesis of how ROS production and removal is tightly linked to stomatal closure in guard cells. The modelling in this chapter also puts forward the hypothesis that ethylene may have a ROS-independent way of producing NO, with effects on pH, Ca^{2+}, and membrane polarity that need to be elucidated.

The model in equations (7.22)–(7.27) predicts that a combined stimulus of 5μM or more of each hormone would result in the arrest of the closure process (Fig. 7.9), a consequence of the fundamental need of increased ROS production for successful closure. This observation requires experimental verification; furthermore, the physiological concentrations of both ABA and ethylene present during environmental stimuli such as bacterial challenge or high humidity that cause stomata to open (Melotto *et al.*, 2006; Zeng *et al.*, 2010) need to be ascertained.

In the natural environment, plants face threats from multiple stimuli; yet under laboratory conditions, mostly single stimuli are studied. This is partly due to the complexity and variability in responses that ensue following exposure to multiple stresses. Using guard cells as a model system we have considered mechanisms for a non-trivial output under a combination of stimuli. This study is a first step towards quantitation of a fundamental

physiological process in plants, which is essential for growth and development.

Chapter 8

Conclusions and further work

8.1 Finishing remarks

In this thesis, we have sought to understand how two signals with overlapping pathways can produce different outputs when applied together than when applied individually. In the case of ABA and ethylene in guard cells this problem has required an interdisciplinary approach, combining traditional experimental biology with computational and mathematical methods. As the understanding about the inherent complexity to the functioning of organisms increases, the need for biologists to liaise with mathematicians and other quantitative scientists becomes even more evident (Yates *et al.*, 2001). The rewards of such interdisciplinary approaches come in the form of a more sophisticated understanding of life. Living organisms are constantly subject to many simultaneous stimuli, and understanding the way in which they sense all these inputs and react to them can be intractable using exclusively experimental or theoretical methods. Thus, we have combined and developed mathematical and computational approaches with experimental methods to investigate stomatal closure under single and combined stimuli. Evidently this work is not the final word in stomatal closure research; on the contrary, (as seen in Chapter 7) our results pose new questions and suggest new avenues for future research. This work is an advance in the quest to understand multiple stimuli in guard cells and other cellular systems.

This work accentuates the importance of transient and dynamical behaviour in cellular processes, as noted previously by Strelkowa and Barahona (2011), and stresses the

fundamental importance of time-course data (as opposed to a single measurement after treatment) as an aid to understanding responses to multiple stimuli. Nonlinear interactions are the norm in biology, and models representing them are capable of a variety of distinct behaviours that depend on the values of their parameters (Kuznetsov, 1998). Time-course data are essential to discriminate between plausible and implausible behaviour; this can be done by narrowing ranges of the parameters or determining their value, aided by methods such as the evolutionary Monte Carlo optimisation algorithm introduced in Chapter 4. Acknowledging the inevitable complexity of large biological models, we have sought to develop methods to reduce models so they can become more tractable without losing the essential features of the system. In Chapter 5 we have made progress in the analysis of activation cascades, their use in model reduction, and as a viable alternative to delay differential equations. Finally, in Chapters 6 and 7 we have brought together experimental observations, mathematical modelling techniques, the results of activation cascades and our new parameter fitting method to study the absence of stomatal closure under simultaneous ABA and ethylene treatment. The modelling process and experimental observations stress the importance of ROS in stomatal closure not only during the initial moments after signal perception, but over the first 60 minutes of signalling and highlight how crucial it is to understand guard cell antioxidant mechanisms, which is another contribution of this work.

8.2 Future work

8.2.1 Evolutionary Monte Carlo methods

Further work on the *Squeeze-and-breathe* method introduced in Chapter 4 shall focus on three main points:

- Establishing a probabilistic framework to allow the *Squeeze-and-breathe* to be used with model selection methods. Currently, Bayesian model selection methods such as ABC-SMC (Toni *et al.*, 2009) require the probability of the posterior to be non-negative, given a prior. The local optimisation step in Algorithm 4.1 can (and does) explore regions outside the prior. A way of overcoming this limitation is to find an acceptable way to assess the probability that the algorithm explores regions outside

the prior or, as mentioned in Chapter 4, to use other distributions such as log-normal or exponential instead of the uniform distribution. The development of alternative model-selection criteria can also be explored. Additionally the local optimisation step of the method could be constrained to find minima only within the prior region, and progressively expanding the prior as needed. Yet another way can be, once the sequence of posteriors becomes nested, to *turn off* the local optimisation step to perform Bayesian analysis in the following iterations.

- It is also important to make the algorithm available to a wider range of optimisation problems, and not only parameter fitting. Presently, the feasible region of the optimisation problem is $\mathbb{R}_+^N$, and in many constrained optimisation problems feasible regions are complicated and even disjoint; algorithms often have to explore infeasible regions before convergence to the minimum (Nocedal and Wright, 2006). Some optimisation methods use *penalty functions* to gauge the gains in terms of reduction of the objective function that are to be obtained from exploring infeasible regions, in the hope that it helps to find a feasible minimum. This feature can be incorporated into the *Squeeze-and-breathe* routine by biasing the posteriors towards feasible regions or ranking local minima according to an established criterion, eg SRES (Runarsson and Yao, 2000, 2005).

- Development and release of a software package for optimisation and parameter fitting, or integration into other currently available software packages.

8.2.2 Activation cascades

The results on activation cascades presented in Chapter 5 so far focus on linear cascades without feedback, Appendix C.5 offers preliminary results on weakly-activated cascades with nonlinear negative feedback. Future work on this subject should include the characterisation of (nonlinear) strongly-activated cascades and cascades with feedback. Including negative feedback is a particularly challenging task because such cascades may display stable steady states and oscillations; for these cascades it may be more convenient to study different behaviours separately, deriving approximations for each case based in

particular regions of the parameter space. Another interesting avenue for further research is to consider the degradation rates of the components of the cascades as random variables, and to derive conditions for optimality analogous to the ones obtained by Chaves *et al.* (2004) for constant rates. Additionally, activation cascades can be studied from a stochastic perspective (eg Csikász-Nagy *et al.* (2010)).

8.2.3 Stomatal closure

The experimental and theoretical investigations of stomatal closure under single and combined ABA and ethylene stimuli have brought about new and interesting questions about the functioning of guard cells. These possibilities for future work share the need for continued interdisciplinary approaches, as has been done in this thesis. Some of the most important tasks for the future are:

- Gathering more time-course data of NO, pH, K^+, Cl^-, under single and combined ABA and ethylene stimuli, for a variety of treatments. These measurements are essential to characterise the behaviour of guard cells and to aid the development and improvement of future models.

- Understanding the ethylene pathway: as mentioned in Sec. 2.4, many components of the ethylene pathway have not yet been uncovered. The interactions directly below ethylene perception to this day remain unknown. Additionally, pH change, NO production, and membrane depolarisation need to be confirmed after treatment with ethylene.

- The hypothesis that the lack of closure following treatment with combined doses of ABA and ethylene was put forward in Chapter 7; this hypothesis needs to be verified experimentally. One way of testing the hypothesis is using mutants. Section 2.2.3 mentions enzymatic antioxidants ascorbate peroxidase, superoxide dismutase, and catalase as part of the antioxidant repertoire of guard cells; plants with the pertinent genes knocked-out can be used to test which of them are involved in controlling ROS during stomatal closure. The use of mutant plants has the risk that other genes may be expressed to compensate for the ones that have been knocked out; this could be

avoided by performing experiments on wild-type plants treated with inactivators or scavengers of antioxidants.

- If it is confirmed that the cross talk of ABA and ethylene occurs in the early stages of signalling, detailed models such as the one in equations (7.1)-(7.8), can be explored and expanded to achieve a full understanding of ABA and ethylene-induced ROS production.

- A stochastic or a hybrid model could be needed to understand aspects of the signalling process. For example if the number of molecules of a particular component (eg a receptor or a protein cluster) is small, then it can be worthwhile to have a stochastic description of its activity, integrated to a larger model containing ODEs. Biophysical approaches may also be of use to understand other aspects of the closure process, eg the biomechanics of stomatal closure.

In addition to the list above, it must be stressed that this work has been exclusively concerned with stomatal closure; as mentioned before, this is a separate process from inhibition of stomatal opening. The transition from one process to the next has not been investigated. An exciting possibility for further research lies in the development of a homeostatic model of guard cells, that incorporates circadian rhythms, stomatal closure, inhibition of stomatal opening, and stomatal aperture. Such a model can be an invaluable tool for understanding how plants react in the short and long term to environmental changes or other pathogens, and may have important applications in the development of more resilient crops and to further understand cellular processes. The approach from this work can be used to understand combined stress signalling pathways where clearly defined outputs and measurements are available, eg cell-death.

Ultimately, profound understanding of stomatal closure and cell signalling under a variety of stimuli should translate into improved knowledge of water and nutrient consumption by plants, whose importance for agricultural applications is difficult to understate.

Appendix A

Competition and antagonism in cellular signals

Here we explore combinations of cell-signalling network topologies and dynamics which exhibit antagonism or competition among incoming signals. This work is inspired from the observation that a two signals such as ABA and ethylene by themselves cause stomatal closure, but do not when they are applied simultaneously (Chapters 2, 6, and 7). Here we show explorations of abstract toy models in which individual signals elicit the same response but in combination display antagonistic behaviour or compete for a limited resource, producing different outcomes together than individually. In Sec. A.3 we show how this approach has been used to analyse how two isoforms of a MAPK compete for activation by the same kinase.

A.1 Antagonistic activation

Suppose that in a cell-signalling system two signals, S_1 and S_2, activate receptors R_1 and R_2, which in turn activate or produce (independently of each other) an output signal F. Here we study the case in which both signals mutually antagonise each other so that S_1 inactivates R_2 and S_2 inactivates R_1. Figure A.1 shows a cartoon of these interactions.

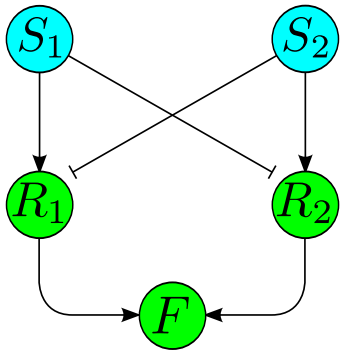

Figure A.1: Antagonistic action of signals S_1 and S_2. The signal S_1 activates R_1, which produces F, and inactivates R_2. In turn, S_2 activates R_2, which also produces F, and inactivates R_1.

A.1.1 Mass-action model

We can represent the interactions from Fig. A.1 (in a simple nondimensional way) using mass-action kinetics. The ODE model of such a signalling network is:

$$\begin{aligned}\frac{\mathrm{d}R_1}{\mathrm{d}t} &= S_1 - (1+S_2)R_1,\\ \frac{\mathrm{d}R_2}{\mathrm{d}t} &= S_2 - (1+S_1)R_2,\\ \frac{\mathrm{d}F}{\mathrm{d}t} &= R_1 + R_2 - F,\end{aligned}$$

with initial conditions $R_1(0) = R_2(0) = F(0) = 0$. The model is linear and can be solved analytically:

$$\begin{aligned}R_1(t) &= \frac{S_1}{1+S_2}\left[1 - e^{-(S_2+1)t}\right],\\ R_2(t) &= \frac{S_2}{1+S_1}\left[1 - e^{-(S_1+1)t}\right],\\ F(t) &= \frac{S_1}{1+S_2}\left[1 + \frac{e^{-(S_2+1)t}}{S_2}\right] + \frac{S_2}{1+S_1}\left[1 + \frac{e^{-(S_1+1)t}}{S_1}\right] - \frac{{S_1}^2 + {S_2}^2}{S_1 S_2}e^{-t}.\end{aligned}$$

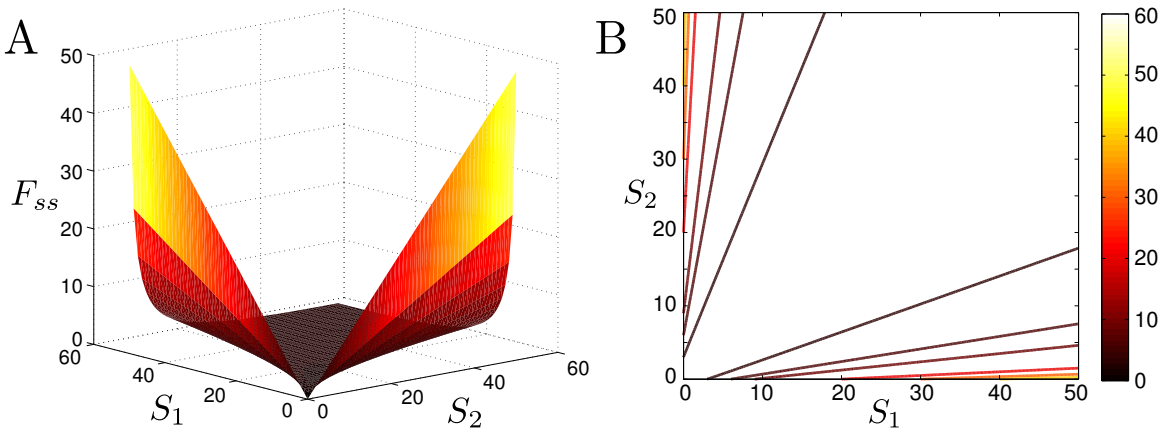

Figure A.2: **A**: Values of F_{ss} given by equation (A.3) under different dose combinations of S_1 and S_2. **B**: Contour map of F_{ss}.

The steady-state solution of the system is:

$$R_{1_{ss}} = \frac{S_1}{S_2 + 1}, \tag{A.1}$$

$$R_{2_{ss}} = \frac{S_2}{S_1 + 1}, \tag{A.2}$$

$$F_{ss} = \frac{S_1}{S_2 + 1} + \frac{S_2}{S_1 + 1}. \tag{A.3}$$

Figure A.2 shows the steady-state behaviour of F_{ss} in equation (A.3). Figure A.2A shows the steady-state landscape of F_{ss} as a function of S_1 and S_2; Fig. A.2B shows the contours of F_{ss}.

If the antagonism is the result of impeding activation rather than actively scavenging the other signal, we have the following equations:

$$\begin{aligned}
\frac{\mathrm{d}R_1}{\mathrm{d}t} &= \frac{S_1}{1 + S_2} - R_1, \\
\frac{\mathrm{d}R_2}{\mathrm{d}t} &= \frac{S_2}{1 + S_1} - R_2, \\
\frac{\mathrm{d}F}{\mathrm{d}t} &= R_1 + R_2 - F.
\end{aligned}$$

The steady state behaviour of this system is the same as in equations (A.1), (A.2), and (A.3).

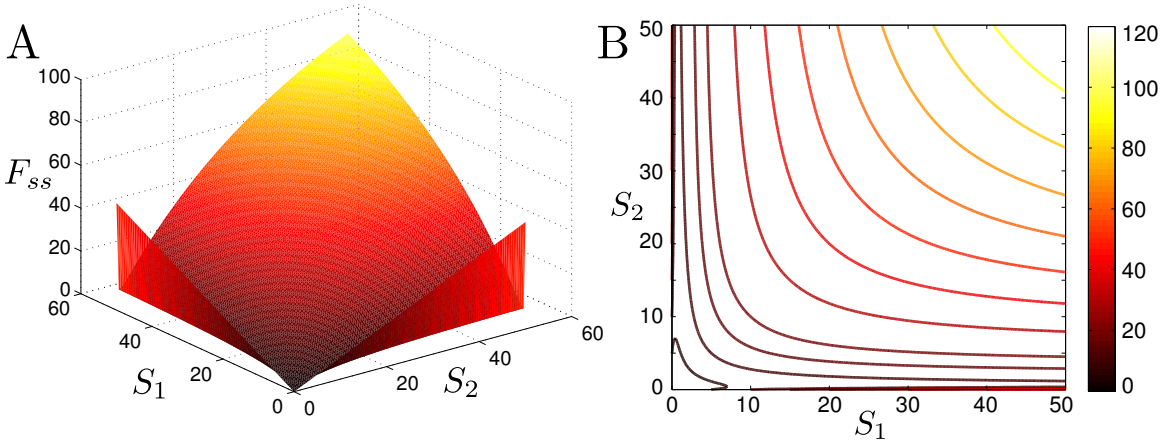

Figure A.3: **A**: Values of F_{ss} given by equation (A.6) under different dose combinations of S_1 and S_2. **B**: Contour map of F_{ss}.

A.1.2 Michaelis-Menten model

If we use Michaelis-Menten kinetics to model the system in Fig. A.1, the ODE model is:

$$\begin{aligned}
\frac{\mathrm{d}R_1}{\mathrm{d}t} &= \frac{S_1}{1+S_1} - \left[1 + \frac{S_2}{1+S_2}\right]\frac{R_1}{1+R_1},\\
\frac{\mathrm{d}R_2}{\mathrm{d}t} &= \frac{S_2}{1+S_2} - \left[1 + \frac{S_1}{1+S_1}\right]\frac{R_2}{1+R_2},\\
\frac{\mathrm{d}F}{\mathrm{d}t} &= \frac{R_1}{1+R_1} + \frac{R_2}{1+R_2} - \frac{F}{1+F},
\end{aligned}$$

with the initial conditions $R_1(0) = R_2(0) = F(0) = 0$. The steady state of the system is:

$$R_{1_{ss}} = \frac{S_1(1+S_2)}{1+2S_2+S_1S_2}, \tag{A.4}$$

$$R_{2_{ss}} = \frac{S_2(1+S_1)}{1+2S_1+S_1S_2}, \tag{A.5}$$

$$F_{ss} = \frac{4\,{S_1}^2{S_2}^2 + 5\,{S_1}^2S_2 + 5\,S_1{S_2}^2 + 2\,{S_1}^2 + 4\,S_1S_2 + 2\,{S_2}^2 + S_1 + S_2}{{S_1}^2S_2 + S_1{S_2}^2 + 5\,S_1S_2 + 2\,S_1 + 2\,S_2 + 1}. \tag{A.6}$$

In Fig. A.3 we show the steady-state behaviour of F given by equation (A.6). Each signal alone causes activation of F proportional to its strength and that it takes just a small value of the second signal for F to decrease to almost zero. As more of the other signal is added, the signals actually cooperate and F is greater than either signal individually.

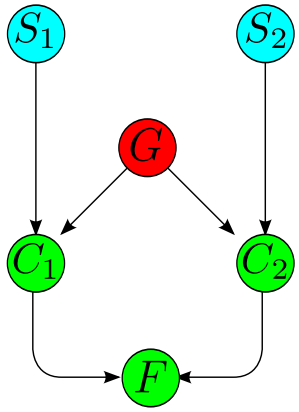

Figure A.4: Competitive activation of two signals S_1 and S_2, both of which need G to produce compounds C_1 and C_2 which produce F.

A.2 Competitive activation

A.2.1 Rate-limiting resource

Two signals S_1 and S_2 combine (or react) with a limited resource G which is needed to produce compounds C_1 and C_2 as shown in Fig. A.4. The compounds can activate an output signal F when a sufficient amount of either one is present. We assume that the total concentration of G is 1 (in arbitrary units) and that

$$G + S_1 \leftrightharpoons C_1,$$
$$G + S_2 \leftrightharpoons C_2.$$

An ODE model for this system is:

$$\frac{\mathrm{d}C_1}{\mathrm{d}t} = S_1\left[1 - C_1 - C_2\right] - C_1,$$
$$\frac{\mathrm{d}C_2}{\mathrm{d}t} = S_2\left[1 - C_1 - C_2\right] - C_2,$$
$$\frac{\mathrm{d}F}{\mathrm{d}t} = \frac{{C_1}^n}{1 + {C_1}^n} + \frac{{C_2}^n}{1 + {C_2}^n} - F,$$

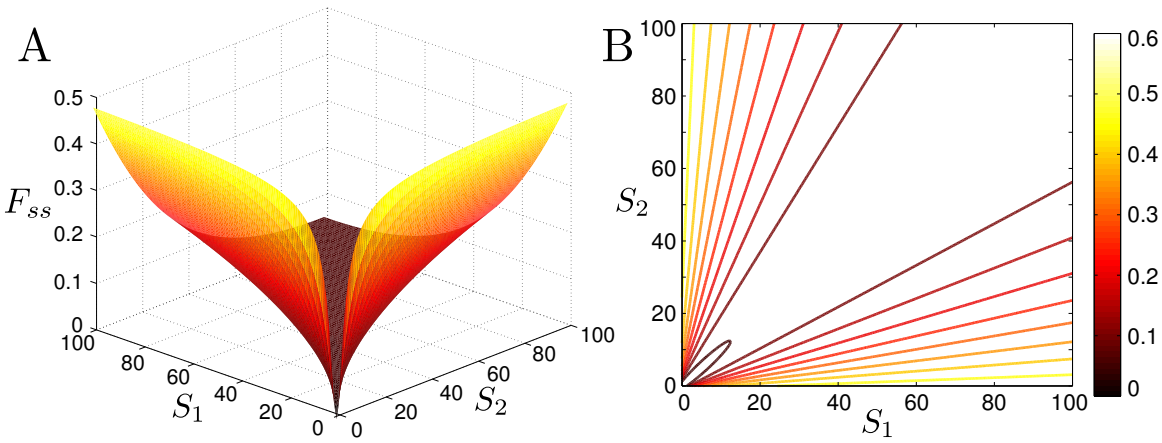

Figure A.5: Response of competitive activation system for $n = 5$. **A**: Values of F_{ss} given by equation (A.9) under different dose combinations of S_1 and S_2. **B**: Contour map of F_{ss}.

where $G(t) = 1-C_1(t)-C_2(t)$. From resting initial conditions ($C_1(0) = C_2(0) = F(0) = 0$) the steady-state of the system is:

$$C_{1_{ss}} = \frac{S_1}{1 + S_2 + S_1}, \tag{A.7}$$

$$C_{2_{ss}} = \frac{S_2}{1 + S_2 + S_1}, \tag{A.8}$$

$$F_{ss} = \frac{\left(\frac{S_1}{1+S_2+S_1}\right)^n}{1 + \left(\frac{S_1}{1+S_2+S_1}\right)^n} + \frac{\left(\frac{S_2}{1+S_2+S_1}\right)^n}{1 + \left(\frac{S_2}{1+S_2+S_1}\right)^n}. \tag{A.9}$$

Figure A.5 shows the response of F_{ss} in equation (A.9) when $n = 5$ to different combinations of S_1 and S_2. Again, each signal by itself produces a response. When $n = 1$ the signals actually cooperate and the response to a combined stimulus is greater than to a single stimulus (Fig. A.6). As n grows, the antagonism between the signals increases.

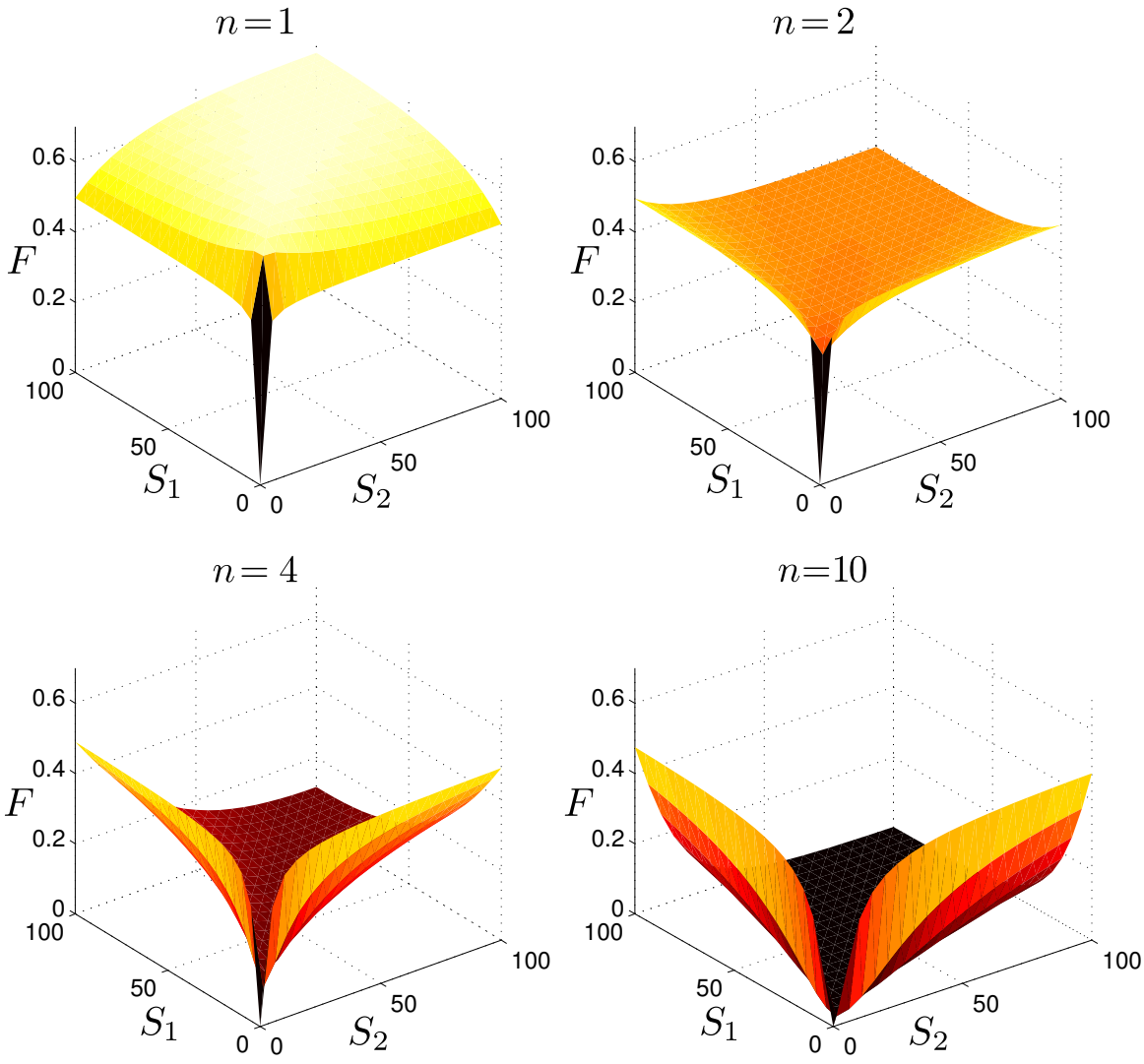

Figure A.6: Response of F_{ss} from equation (A.9) to combinations of S_1 and S_2 and different values of n.

A.2.2 Compound Michaelis-Menten forms

Suppose an enzyme E catalyses the production of a substance P from two different substrates S_1 and S_2 in the following chemical reactions:

$$S_1 + E \underset{k_2}{\overset{k_1}{\rightleftharpoons}} C_1 \xrightarrow{k_3} P + E$$
$$S_2 + E \underset{k_5}{\overset{k_4}{\rightleftharpoons}} C_2 \xrightarrow{k_6} P + E.$$

The enzyme binds to the substrates to form a complex which can either dissociate or catalyse the reaction. If $[E]$, $[S_1]$, $[S_2]$, $[C_1]$, $[C_2]$, and $[P]$ denote the concentrations of the reactants and the products of the reactions, and the substrates are abundant enough (ie the concentrations can be considered to remain constant), then the ODE system describing the reactions is:

$$\frac{\mathrm{d}[E]}{\mathrm{d}t} = -k_1[S_1][E] - k_4[S_2][E] + (k_2 + k_3)[C_1] + (k_5 + k_6)[C_2], \tag{A.10}$$
$$\frac{\mathrm{d}[C_1]}{\mathrm{d}t} = k_1[S_1][E] - (k_2 + k_3)[C_1], \tag{A.11}$$
$$\frac{\mathrm{d}[C_2]}{\mathrm{d}t} = k_4[S_2][E] - (k_5 + k_6)[C_2], \tag{A.12}$$
$$\frac{\mathrm{d}[P]}{\mathrm{d}t} = k_3[C_1] + k_6[C_2]. \tag{A.13}$$

Because $\frac{\mathrm{d}[E]}{\mathrm{d}t} + \frac{\mathrm{d}[C_1]}{\mathrm{d}t} + \frac{\mathrm{d}[C_2]}{\mathrm{d}t} = 0$, we express the total concentration of enzyme as $E_T = [E] + [C_1] + [C_2]$, $E_T \in \mathbb{R}^+$. We eliminate equation (A.10) using this conservation relation. Now equations (A.11) and (A.12) become:

$$\frac{\mathrm{d}[C_1]}{\mathrm{d}t} = k_1[S_1](E_T - [C_1] - [C_2]) - (k_2 + k_3)[C_1], \tag{A.14}$$
$$\frac{\mathrm{d}[C_2]}{\mathrm{d}t} = k_4[S_2](E_T - [C_1] - [C_2]) - (k_5 + k_6)[C_2]. \tag{A.15}$$

The quasi-steady-state approximation (QSSA) states that before any meaningful amounts of P are produced, the enzymes and complexes reach an equilibrium (Murray, 2005; Segel and Slemrod, 1989). Thus, equations (A.14) and (A.15) are both equal to zero, because

they do not depend on equation (A.13) and we can analyse them in isolation:

$$0 = k_1[S_1](E_T - [C_1] - [C_2]) - (k_2 + k_3)[C_1],$$
$$0 = k_4[S_2](E_T - [C_1] - [C_2]) - (k_5 + k_6)[C_2].$$

If $k_a = (k_2 + k_3)/k_1$ and $k_b = (k_5 + k_6)/k_4$ then

$$[C_1] = \frac{k_b E_T[S_1]}{k_b[S_1] + k_a[S_2] + k_a k_b}, \tag{A.16}$$
$$[C_2] = \frac{k_a E_T[S_2]}{k_b[S_1] + k_a[S_2] + k_a k_b}. \tag{A.17}$$

The production of P can be expressed as

$$\frac{\mathrm{d}[P]}{\mathrm{d}t} \approx \frac{E_T(k_3 k_b[S_1] + k_6 k_a[S_2])}{k_b[S_1] + k_a[S_2] + k_a k_b}, \tag{A.18}$$

which is a compound Michaelis-Menten form. When $[S_1] > 0$ and $[S_2] = 0$ the expression in equation (A.16) becomes proportional to the standard Michaelis-Menten form:

$$[C_1] = \frac{E_T k_3[S_1]}{[S_1] + k_a},$$

likewise when $[S_1] = 0$ and $[S_2] > 0$ we have

$$[C_2] = \frac{E_T k_6[S_2]}{[S_2] + k_b}.$$

Figure A.7 shows an example of the compound Michaelis-Menten form (A.18) with $E_T = 1$, $k_3 = 2$, $k_6 = 0.8$, $k_a = 5$, and $k_b = 1$. On Fig. A.7A we show the value of $\frac{\mathrm{d}P}{\mathrm{d}t}$ with different signal combinations, and on Fig. A.7B we show the cases when one of the two signals is zero (bold blue and dashed red lines) and when the two signals are the same (dash-dotted green line). Here we can see how one signal could affect the response of the other.

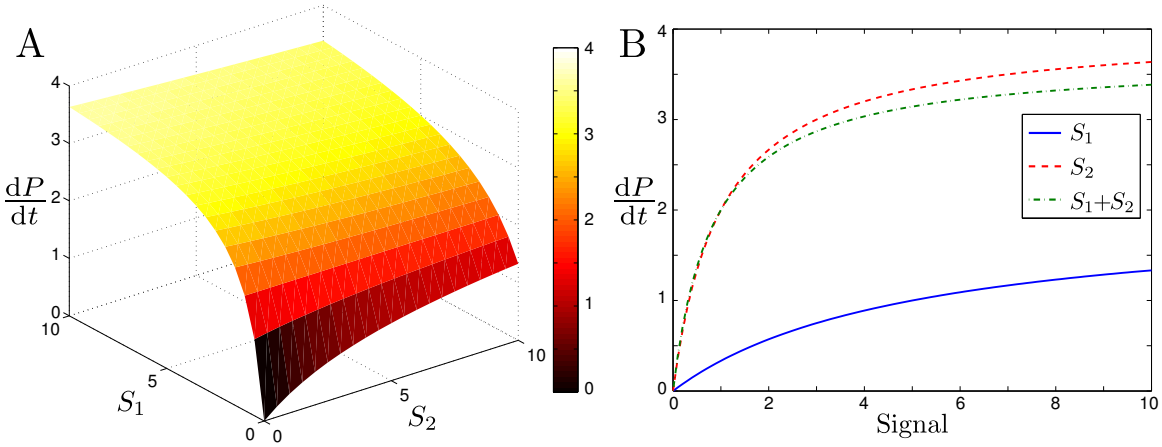

Figure A.7: Example of compound Michaelis-Menten responses of equation (A.18) when $E_T = 1$, $k_3 = 2$, $k_6 = 0.8$, $k_a = 5$, and $k_b = 1$. **A**: Response to combinations of S_1 and S_2. **B**: Response to a single signal (bold blue and dashed red lines) and to $S_1 + S_2$ with "equal doses" of each signal (dash-dotted green line).

A.3 Competition for activation among MAPK isoforms

The approaches to understanding antagonism and competition for limited resources discussed above has been used by Harrington *et al.* (2011) to understand the dynamics of the activation of two MAPKs that share the same activating kinase (MAPKK). The mitogen activated protein kinases (MAPK) are important relays in many of these signal transduction processes (Marks *et al.*, 2009); they are involved in regulating cellular fates such as proliferation, differentiation and apoptosis (Muller and Ram, 2010). The most widely studied MAPKs, Erk1 and Erk2, are activated through phosphorylation by Mek, their MAPK kinase (MAPKK), and Mek in turn is activated by its kinase (MAPKKK) Raf. Activated kinases, such Erk or Mek, are deactivated via dephosphorylation by their respective phosphatases. Erk is encoded by two genes, Erk1 and Erk2 differ only subtly at the sequence level; however, Erk1 and Erk2 appear to have different biological characteristics (Mazzucchelli *et al.*, 2002; Pagès *et al.*, 1999).

We study Erk1/2 activation in two cell types, HeLa and NIH 3T3, and under different conditions. The two Erk isoforms (Erk1 and Erk2) exist in one of three states: inactive (Erk), bound to Mek (Merk), or active (Erk_P). Erk transitions from inactive to active through Mek, free phosphorylated Mek reversibly interacts with inactive Erk. Upon Mek-

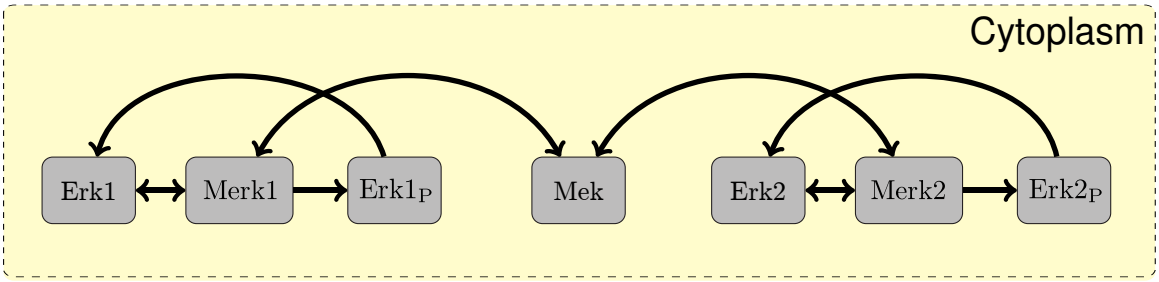

Figure A.8: Schematic of Erk1/2 activation by Mek. Double arrows denote reversible reactions, and single arrows; irreversible reactions. Image modified from Harrington *et al.* (2011).

Reaction	Forward rate (μM^{-1} s^{-1})	Reverse rate (s^{-1})
$Erk1 + Mek \rightleftharpoons Merk1$	8.8×10^{-1}	8.8×10^{-2}
$Merk1 \rightarrow Mek + Erk1_P$	3×10^{-1}	
$Erk2 + Mek \rightleftharpoons Merk2$	8.8×10^{-1}	8.8×10^{-2}
$Merk2 \rightarrow Mek + Erk2_P$	2×10^{-1}	
$Erk1_P \rightarrow Erk1$	1.4×10^{-2}	
$Erk2_P \rightarrow Erk2$	1.4×10^{-2}	

Table A.1: Reactions of Erk1/2 activation by Mek, and deactivation. Forward and, where pertinent, backward reaction rates are included. The reactions are illustrated by Fig. A.8 and described by equations (A.19)-(A.25). All reaction rates take from Fujioka *et al.* (2006).

Erk binding and formation of an intermediary complex (Merk), Mek phosphorylates Erk (Erk$_P$) and then dissociates. In order for Erk$_P$ to revert back to its inactive form, it undergoes dephosphorylation by a phosphatase. Fig. A.8 shows a schematic of Erk1/2 activation by Mek. The mass-action model describing the reactions in Table A.3 are:

$$\frac{\mathrm{d}[Erk1]}{\mathrm{d}t} = -k_1[Erk1][Mek] + k_{-1}[Merk1] + k_9[Erk1_P], \tag{A.19}$$

$$\frac{\mathrm{d}[Mek]}{\mathrm{d}t} = -k_1[Erk1][Mek] - k_5[Erk2][Mek] + (k_{-1} + k_2)[Merk1] + (k_{-5} + k_6)[Merk2], \tag{A.20}$$

$$\frac{\mathrm{d}[Merk1]}{\mathrm{d}t} = k_1[Erk1][Mek] - (k_{-1} + k_2)[Merk1], \tag{A.21}$$

$$\frac{\mathrm{d}[Erk1_P]}{\mathrm{d}t} = k_2[Merk1] - k_9[Erk1_P], \tag{A.22}$$

Parameter	Description
k_1	Forward interaction rate between cytoplasmic Mek and Erk1
k_{-1}	Reverse interaction rate between cytoplasmic Mek and Erk1
k_2	Phosphorylation rate of cytoplasmic Erk1
k_5	Forward interaction rate between cytoplasmic Mek and Erk2
k_{-5}	Reverse interaction rate between cytoplasmic Mek and Erk2
k_6	Phosphorylation rate of cytoplasmic Erk2
k_7	Forward interaction rate between cytoplasmic Mek and Erk2
k_{-7}	Reverse interaction rate between cytoplasmic Mek and Erk2
k_9	Dephosphorylation rate of cytoplamsic Erk1$_P$
k_{11}	Dephosphorylation rate of cytoplasmic Erk2$_P$

Table A.2: Description of the parameters in equations (A.19)-(A.25).

$$\frac{\mathrm{d}[Erk2]}{\mathrm{d}t} = -k_5[Erk2][Mek] + k_5[Merk2] + k_{11}[Erk2_P], \tag{A.23}$$

$$\frac{\mathrm{d}[Merk2]}{\mathrm{d}t} = k_5[Erk2][Mek] - (k_5 + k_6)[Merk2], \tag{A.24}$$

$$\frac{\mathrm{d}[Erk2_P]}{\mathrm{d}t} = k_6[Merk2] - k_{11}[Erk2_P], \tag{A.25}$$

and conservation relations:

$$[Mek_T] = [Merk1] + [Merk2] + [Mek],$$
$$[Erk1_T] = [Erk1] + [Erk1_P] + [Merk1],$$
$$[Erk2_T] = [Erk2] + [Erk2_P] + [Merk2].$$

The parameters of the model are described in Table A.3. We study the system under two different conditions: baseline and limited Mek, in each case the model has different initial conditions. Under baseline conditions we have that $[Erk1]_0 = 0.2$, $[Erk2]_0 = 0.8$, $[Mek]_0 = 1.0$ (Fujioka *et al.*, 2006; Marchi *et al.*, 2008), and all other variables are zero at $t = 0$. We explore the effects that limiting the amount of Mek would have on the system by setting $[Mek]_0 = 0.2$, and studying the system in parallel to baseline conditions.

We investigate the effects of the phosphorylation/dephosphorylation rates on the relative abundances of Erk2$_P$ to Erk1$_P$ at steady state in both the baseline Mek and the limited Mek scenarios. These parameters have been identified by Harrington *et al.* (2011)

as the most sensitive in the model. The log-scale indicator

$$\varphi = \log\left(\frac{[Erk2_P]}{[Erk1_P]}\right),$$

can identify the dominance of a specific phospho-Erk isoforms by orders of magnitude as physical conditions are changed. For example, a value of 1.0 indicates ten times more Erk2$_P$ than Erk1$_P$ whereas a value of -1.0 establishes there is ten times more Erk1$_P$ than Erk2$_P$. The phosphorylation rates and dephosphorylation rates are each varied separately from 10^{-3} to 10^{1} in line with biophysical considerations and φ is evaluated (Fig. A.9A,B). As the rate of phosphorylation of Erk1 (k_2) increases, values of φ become negative indicating dominance by Erk1$_P$ (Fig. A.9A (ii)), whereas an increase in the Erk2 phosphorylation rate (k_6) demonstrates $Erk2$ prevalence. The effect of the phosphorylation rate on φ is asymmetric, meaning that there is a larger region in the parameter space of values giving rise to Erk2$_P$ dominance (see non-overlapping curves in Fig. A.9A (i)). Under limited Mek conditions the maximum/minimum values of φ are larger and smaller, respectively than the baseline conditions. An increase in dephosphorylation rate of Erk1 (k_9) results in a larger positive φ, or an Erk2$_P$ steady-state bias, and vice-versa for Erk2 dephosphorylation (k_{11}); this is also reflected in the heat map asymmetry, with a bias towards Erk2$_P$, which is more easily activated than Erk1$_P$. This asymmetry is also apparent in the time course (Fig. A.9B (i)). Under limited Mek conditions, the phosphorylation/dephosphorylation rate parameter space has a larger region for possible competition scenarios (see largest (white) and smallest (black) values of φ, corresponding to high Erk2$_P$ and Erk1$_P$ dominance in Fig. A.9A,B (iii)). Such dominance suggests that for certain phosphorylation/dephosphorylation rates, a limited stimulus would more strongly favour an Erk1$_P$ response than non-limited stimuli. Figure A.9C illustrates how the total amount of Erk may vary across cells affects activation states. For small total Mek, as well as baseline conditions, the initial amount greatly affects the steady-state value and it gives the expected result that as total Erk1 in the system increases, the total Erk1$_P$ also increases (Fig. A.9C). Unlike the phosphorylation/dephosphorylation cases, the value of φ at a given total $([Erk1_T], [Erk2_T])$ point does not change as Mek becomes limited (heat map indicator colours are nearly identical).

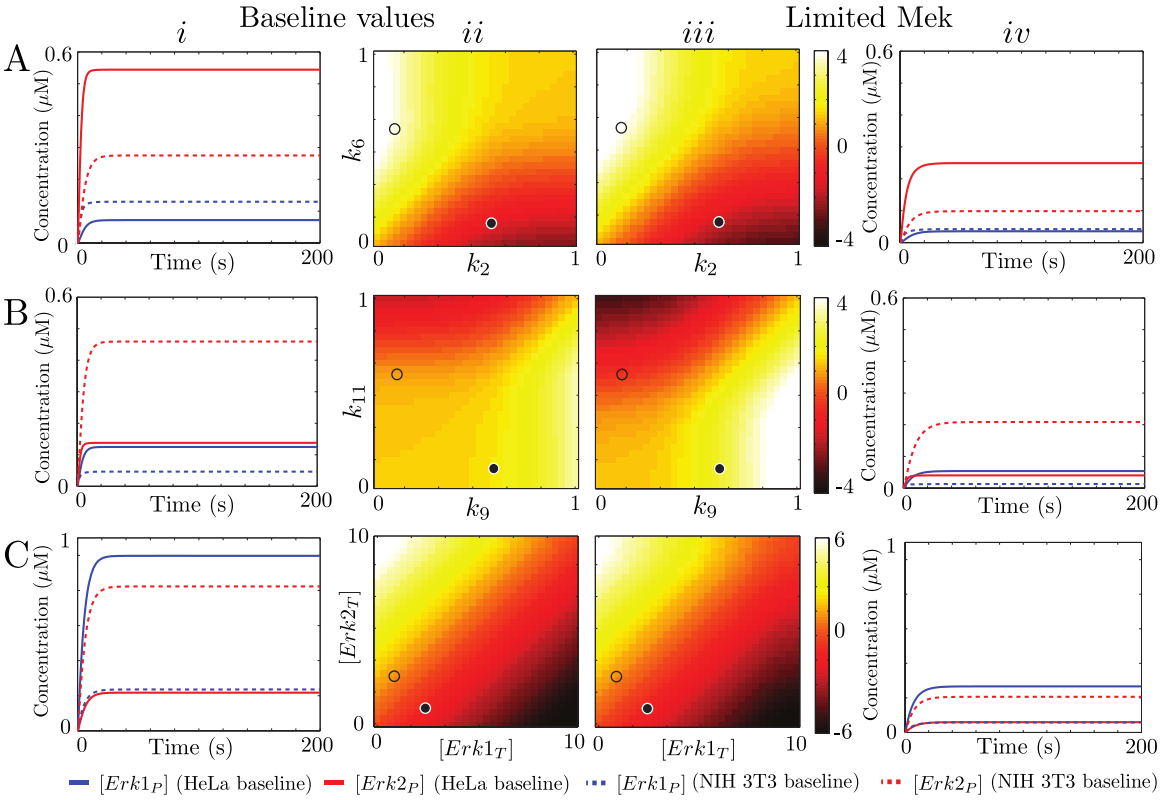

Figure A.9: Competition of Erk1/2 for activation. Competition in minimal model at baseline ($[Mek]_0 = 0.8$) and limited ($[Mek]_0 = 0.2$) conditions. Competition scenarios are studied using heat maps and time courses. Heat maps (ii-iii) of parameter /initial condition values are given along the horizontal axis and vertical axis are varied and the associated colour at each parameter/initial condition combination correspond to $\varphi = \log([Erk2_P]/[Erk1_P])$. Solid circles and open circles in heat maps are shown in time course of $[Erk1_P]$ (blue) and $[Erk2_P]$ (red) in (i, iv) where a solid circle corresponds to a solid line and an open circle corresponds to a dotted line. **A**: Effects of phosphorylation rates on Erk1/2p. Values of (k_2, k_6) are varied, solid circle is high $k_2 = 0.6$, low $k_6 = 0.1$ and open circle is low $k_2 = 0.1$, high $k_6 = 0.6$. **B**: Effects of dephosphorylation rates on $[Erk1_P]$ and $[Erk2_P]$. Values of (k_9, k_{10}) are varied, solid circle is high $k_9 = 0.6$, low $k_{10} = 0.1$ and open circle is low $k_9 = 0.1$, high $k_{10} = 0.6$. **C**: Effects of $[Erk1]_0$ and $[Erk2]_0$. Values of total Erk concentrations $([Erk1_T], [Erk2_T])$ are varied, solid circle is high $[Erk1_T] = 2.0$, low $[Erk2_T] = 0.5$, and open circle is low $[Erk1_T] = 0.5$, high $[Erk2_T] = 2.0$. Image modified from Harrington *et al.* (2011).

The model (A.19)-(A.25) set in a competition framework provides a number of insights, specifically it highlights that phosphorylation/dephosphorylation rates play an important role in the steady-state behaviour of Erk1_P and Erk2_P. Under limited Mek conditions, the parameter space exploration suggests there is a strong effect (large $|\varphi|$) on the response. The value of the initial conditions can induce a Erk2_P or Erk1_P dominated response and limiting Mek alters the steady-state value of this response. More generally, this model provides a simple framework for gaining insight into the components which control the competition between Erk1 and Erk2 for its kinase Mek, and we provide a indicator φ for giving Erk1_P or Erk2_P cell response. See Harrington *et al.* (2011) for more details.

Appendix B

Data fitting tables

B.1 Bliss-Painter-Marr model data

t	R
0	0
20	43.5373
40	13.3667
60	140.8903
80	29.2816
100	108.1722
120	19.0093
140	75.0065
160	14.4018
180	50.4473
200	217.1082

Table B.1: BPM data.

Table B.1 shows data obtained from a simulation of the BPM model from equations (4.10) using parameters $\alpha = 240$ and $\beta = 0.15$, initial conditions $R(0) = 0$, $E(0) = 0$, $P(0) = 0$, and adding random noise sampled from a $N(0, 15^2)$ distribution. Only the data for variable R was obtained.

t	S	I	R
0.6	0.12	13.17	9.42
1.0	0.12	7.17	11.19
2.0	0.10	2.36	10.04
3.0	0.38	0.92	6.87
4.0	1.00	0.62	4.45
5.0	1.20	0.17	3.01
6.0	1.46	0.28	1.76
7.0	1.38	0.10	1.29
8.0	1.57	0.03	0.82
9.0	1.46	0.29	0.52
10.0	1.25	0.10	0.23
11.0	1.56	0.22	0.20

Table B.2: SIR data.

B.2 Susceptible-Infected-Recovered model data

Table B.2 shows data for the SIR model generated from equations (4.11) using initial conditions $S(0) = 20$, $I(0) = 10$, and $R(0) = 0$ with added random noise sampled from a $N(0, 0.2^2)$ distribution as appears in Toni *et al.* (2009).

B.3 Repressilator data

Table B.3 contains the simulated repressilator data corresponding to the three mRNA variables in model (4.13). The data were generated by Toni and Stumpf (2010) using parameters $\alpha_0 = 1$, $n = 2$, $\beta = 5$, and $\alpha = 1000$, starting from initial conditions $m_1(0) = 0$, $p_1(0) = 2$, $m_2(0) = 0$, $p_2(0) = 1$, $m_3(0) = 0$, and $p_3(0) = 3$, with random noise added from a $\mathcal{N}(0, 5^2)$ distribution.

B.4 Genetic switch data

Tables B.4 and B.5 show the fluorescent response of IPTG-induced genetic switches described in Wang (2010); Wang *et al.* (2011).

t	m_1	m_2	m_3
0	0	0	0
0.6	2.04	28.99	20.96
4.2	32.19	11.29	7.49
6.2	4.13	10.61	44.25
8.6	2.15	55.27	7.12
13.4	5.09	9.49	60.52
16	1.07	68.56	8.1
21.4	3.67	10.62	63.76
27.6	39.01	-1.95	22.9
34.4	73.83	3.53	6.27
39.8	8.54	63.87	10.59
40.6	17.62	39.68	6.5
45.2	11.96	-0.6	70.56

Table B.3: Repressilator data.

t	0mM	0.0004mM	0.0016mM	0.0063mM	0.025mM	0.1mM	0.4mM	1.6mM	6.4mM	12.8mM
0	0	0	0	0	0	0	0	0	0	0
140	88.6	177.8	174.4	197.8	210.4	1043.6	3945.8	5971	6643.8	6521.8
160	120.2	156.4	160.6	165.6	209.8	1300.8	4695.2	6768.4	7361.8	7513.8
180	66.6	96.4	94.6	126.4	171.6	1438.4	5238.8	7465.2	7801	8002.4
200	42.8	72.2	76.2	88	134.2	1578	5658	7914	8458	8542.8
220	37	64.8	61.2	55	135.8	1667	5799.6	8380.2	8976	8914.8
240	39.6	56.6	60.4	65.8	142.8	1758.6	6108.6	8601.4	9172.6	8957
260	36.2	47.6	62	69.8	143.6	1859.8	6104	9041.8	9528.6	9252.8
280	50.8	55.6	58.2	74.2	170.6	1968.2	6554.4	9071.6	9449	9018.4
300	39.6	51	40.8	60.2	197.8	2143.4	6452.2	8396.2	9269.2	9261.2
320	50.4	62.8	65.6	82	273.6	2317.8	6880.8	8941.2	9887.6	9982.8
340	53.8	71.4	71	88.6	296	2512.8	7052.2	8972.8	9694.6	10108
360	45.6	66	61.6	69.2	340.8	2639.2	7047.8	9103.6	9911	10018.4

t	0mM	0.0004mM	0.0016mM	0.0063mM	0.025mM	0.1mM	0.4mM	1.6mM	6.4mM	12.8mM
0	0	0	0	0	0	0	0	0	0	0
140	215	163.4	124.8	134	119	230.4	721.2	1001.8	1095.8	701
160	141.6	116.6	95.4	86	40	320.6	937	1112.2	1054	903.2
180	131.6	112.2	117.6	84	81	252.2	825.2	727.4	1026.8	679.2
200	69.8	42.4	37.8	39	44.2	225.2	688.4	829.8	761.6	584.6
220	55	58.4	59	60.6	50.4	169.2	645.8	713.6	739.6	454
240	38.8	48	30.8	43.4	42.2	148.8	366	418.6	453.8	668.2
260	42.2	44	48.6	41	53.8	152.8	496.4	638.4	547.8	626.2
280	55.2	54.4	51.8	53.6	76	257.2	498.2	722.2	889.8	606.2
300	50.4	57.4	62	67.8	95	339.8	447.4	835.6	693.2	602.6
320	52.6	69.6	78.4	81.2	146.8	385.8	540.4	776.4	1084.2	580
340	57	60.6	73.8	65.6	144.6	401.2	466.4	396.6	560.4	702
360	61.6	73.2	77.2	68.6	151	400	374.8	251	742	436.2

Table B.4: *gfp*30 fluorescence measurements (top) and standard deviations (bottom).

t	0mM	0.0004mM	0.0016mM	0.0063mM	0.025mM	0.1mM	0.4mM	1.6mM	6.4mM	12.8mM
0	0	0	0	0	0	0	0	0	0	0
140	149.1	199.7	107.4	124.6	242.4	801.9	2682.7	4292.3	4633.3	4923.8
160	96	212.2	121.6	78.4	199.3	945	3192.9	4893.7	5243.3	5572.6
180	64.3	178.7	73.7	40.4	158.7	1083.8	3598.4	5362.7	5762.6	6139.4
200	32.2	92.5	43.2	43.7	135.1	1190.5	3961.4	5901.6	6282.9	6499.9
220	56.4	86.5	51.5	43.5	142.8	1320.4	4274.4	6218.6	6589.5	6866.5
240	42.4	54.6	16.5	23.9	116.3	1330.6	4424.9	6247.9	6514.3	6815.1
260	31	49.9	11.3	13.4	100.4	1422.8	4583.5	6531	6917.5	7177.6
280	34.7	55.5	13	16.4	107.1	1535.8	4680.4	6609.6	7247.2	7290.1
300	33.2	46.1	21.7	22.1	129.7	1675.5	4958.5	6949.3	7620.3	7631.3
320	29.5	39	8.7	22.5	154	1824.5	5122.3	7053.4	7642.7	7645.1
340	31.2	43.2	19.1	27.1	172.2	1836.2	5282.7	7156.9	7661.2	7889.3
360	28	40	10.9	28.9	202.4	1979.3	5456.4	7245.6	7899.1	7910.6

t	0mM	0.0004mM	0.0016mM	0.0063mM	0.025mM	0.1mM	0.4mM	1.6mM	6.4mM	12.8mM
0	0	0	0	0	0	0	0	0	0	0
140	89.4	85.8	209.2	120.8	77.8	175.8	383.6	295.4	332.6	382
160	59.4	23	166.4	111.6	40.6	188.8	572.2	391.6	430.6	326.2
180	31.6	38.6	135.4	51.2	24.8	210.6	597	370.8	467.6	363.8
200	45.2	60.4	83.2	65.2	42	166	573.2	273.6	341.6	337
220	14	27.4	90.2	51.2	25	90	513.8	249.6	234	145.2
240	25.2	32.2	53.8	30.6	16.2	70	475.2	187.6	464.8	168
260	14.8	17.2	47.4	23.8	14.2	68.8	511.8	256	300.6	214
280	20	15.4	46.6	16.6	15.8	70.6	395.8	237.6	313.6	454.6
300	17.8	17.8	37.8	29.8	29.2	178.2	486.6	383.8	416.2	377.2
320	21	21.2	43	26.4	46.8	216.2	519.6	507.4	674.8	227
340	26	22.2	36.8	25.4	46.6	340.8	495.6	655.6	594.2	299.4
360	15.2	13	38.4	8.6	50	350.4	604.8	434.2	853.8	387.8

Table B.5: *gfp*34 fluorescence measurements (top) and standard deviations (bottom).

Appendix C

Analysis of activation cascades

C.1 Calculation of the output of the linear cascade, $x_n(t)$

In this section we introduce necessary notation to present the calculations for obtaining the solutions to equation (5.3).

C.1.1 The Laplace transform

If $f(t)$ is a function integrable over $t \geq 0$ its Laplace transform is given by (Kreyszig, 2006):

$$F(s) = \mathscr{L}(f) = \int_0^\infty e^{-st} f(t) \mathrm{d}t.$$

The inverse of the Laplace transform is given by

$$f(t) = \mathscr{L}^{-1}(F) = \frac{1}{2\pi i} \lim_{T \to \infty} \int_{\gamma - iT}^{\gamma + iT} e^{st} F(s) \mathrm{d}s,$$

where γ is large enough so that the line of integration is beyond all singularities of F. The Laplace transform (and its inverse) are linear so that

$$\mathscr{L}(af(t) + g(t)) = a\mathscr{L}(f) + \mathscr{L}(g),$$

where a is a constant. An important property of the Laplace transform is that

$$\mathscr{L}(f') = s\mathscr{L}(f) - f(0),$$

where $f' = \frac{\mathrm{d}f}{\mathrm{d}t}$. This can be applied successively to n-th order derivatives. The convolution of functions f and g is

$$(f * g) = \int_0^t f(\tau)g(t-\tau)\mathrm{d}\tau = \int_0^t f(t-\tau)g(\tau)\mathrm{d}\tau,$$

and

$$\mathscr{L}(f * g) = \mathscr{L}(f)\mathscr{L}(g).$$

C.1.2 The incomplete gamma function: definitions and notation

The gamma function is defined as (Abramowitz and Stegun, 1964):

$$\Gamma(a) = \int_0^\infty e^{-s}s^{a-1}\mathrm{d}s, \qquad \mathrm{Re}(a) > 0. \tag{C.1}$$

One well-known property of the gamma function is $\Gamma(n) = (n-1)!$, $n \in \mathbb{N}$.

The *lower incomplete* gamma function is given by:

$$\gamma(a,t) = \int_0^t e^{-s}s^{a-1}\mathrm{d}s, \qquad \mathrm{Re}(a) > 0 \tag{C.2}$$

A different way of writing $\gamma(a,t)$ when $a = n \in \mathbb{N}$ is (Paris, 2010):

$$\gamma(n,t) = (n-1)!\left(1 - e^{-t}\sum_{k=0}^{n-1}\frac{t^k}{k!}\right). \tag{C.3}$$

The *normalised* lower incomplete gamma function, which we use in our calculations, is defined as:

$$\mathrm{P}(a,t) = \frac{\gamma(a,t)}{\Gamma(a)}. \tag{C.4}$$

C.1.3 Constant stimulus: derivation of Eq. (5.12)

The solution of the ODE system (5.3) for $\mathbf{x}(0) = 0$ when $R(t) = \alpha_1$ is given in Eq. (5.11):

$$\mathbf{x}(t) = \alpha_1 \mathbf{A}^{-1} \left[e^{t\mathbf{A}} - \mathbf{I}_n \right] \mathbf{e}_1. \tag{C.5}$$

When $\mathbf{A} = \tilde{\mathbf{A}}$ (defined in equation (5.10)), it can be shown by mathematical induction that the n-th component of the solution is:

$$x_n(t) = \left(\frac{\alpha_{(n)}}{\beta} \right)^n \left(1 - e^{-\beta t} \sum_{k=0}^{n-1} \frac{(\beta t)^k}{k!} \right). \tag{C.6}$$

A similar result is obtained by Lucius *et al.* (2003) from the analysis of linear models of n-step DNA unwinding.

Using (C.3) and the properties of the gamma function, equation (C.6) becomes:

$$x_n(t) = \left(\frac{\alpha_{(n)}}{\beta} \right)^n \frac{\gamma(n, \beta t)}{\Gamma(n)} = \left(\frac{\alpha_{(n)}}{\beta} \right)^n \mathrm{P}(n, \beta t), \tag{C.7}$$

which is the expression presented in equation (5.12).

C.1.4 Exponentially decaying stimulus: derivation of Eq. (5.14)

The solution of the ODE system (5.3) for $\mathbf{x}(0) = 0$ when $R(t) = \alpha_1 e^{-\lambda t}$ is given in (5.13):

$$\mathbf{x}(t) = \alpha_1 \left[e^{t\mathbf{A}} - e^{-\lambda t} \mathbf{I}_n \right] \mathbf{A}^{-1} \left[\mathbf{I}_n + \lambda \mathbf{A}^{-1} \right]^{-1} \mathbf{e}_1. \tag{C.8}$$

When $\mathbf{A} = \tilde{\mathbf{A}}$ and $\beta \neq \lambda$, we can use mathematical induction to show that

$$x_n(t) = \left(\frac{\alpha_{(n)}}{\beta - \lambda} \right)^n \left(e^{-\lambda t} - e^{-\beta t} \sum_{k=0}^{n-1} \frac{t^k (\beta - \lambda)^k}{k!} \right) \tag{C.9}$$

$$= \left(\frac{\alpha_{(n)}}{\beta - \lambda} \right)^n e^{-\lambda t} \frac{\gamma(n, (\beta - \lambda)t)}{\Gamma(n)} = \left(\frac{\alpha_{(n)}}{\beta - \lambda} \right)^n e^{-\lambda t} \mathrm{P}(n, (\beta - \lambda)t), \tag{C.10}$$

and we obtain the top expression of equation (5.14). When $\beta = \lambda$, we solve sequentially the ODEs (5.3) and use mathematical induction to get

$$x_n(t) = \frac{\left(\alpha_{(n)}t\right)^n e^{-\beta t}}{\Gamma(n+1)}, \tag{C.11}$$

the bottom expression of (5.14).

C.1.5 Sinusoidal stimulus: derivation of Eq. (5.16)

To obtain an explicit solution for $x_n(t)$ when the input to equation (5.3) is

$$R(t) = \alpha_1(1 + \sin(\omega t)),$$

we use the property of the Laplace transform $\mathscr{L}(f') = s\mathscr{L}(f) - f(0)$ to obtain the transform of the equations of the system:

$$\mathscr{L}(x_1) = \hat{x}_1(s) = \frac{\alpha_1\omega}{(s^2+\omega^2)(\beta+s)} + \frac{\alpha_1}{s(\beta+s)},$$

where $\mathscr{L}(x_1) = \hat{x}_1(s)$ and $x_1(0) = 0$. It can be shown by mathematical induction that

$$\hat{x}_n(s) = \frac{\alpha_{(n)}^n}{s(\beta+s)^n}\frac{\alpha_{(n)}^n\omega}{(s^2+\omega^2)(\beta+s)^n}. \tag{C.12}$$

To obtain $x_n(t)$, we must compute the inverse transform of $\hat{x}_n(s)$. First we obtain the inverse transform of the first term of equation (C.12); we use the properties of the Laplace transform of convolutions to find two $f(t)$ and $g(t)$ such that

$$\mathscr{L}(f*g) = \mathscr{L}(f)\mathscr{L}(g) = \frac{\alpha_{(n)}^n}{s(\beta+s)^n}$$

where

$$(f*g)(t) = \int_0^t f(\tau)g(t-\tau)\mathrm{d}\tau = \int_0^t g(\tau)f(t-\tau)\mathrm{d}\tau.$$

Note that (Kreyszig, 2006):

$$f(t) = \mathscr{L}^{-1}\left(\frac{1}{(\beta+s)^n}\right) = \frac{1}{\Gamma(n)}t^{n-1}e^{-\beta t}, \qquad g(t) = \mathscr{L}^{-1}\left(\frac{1}{s}\right) = 1,$$

and

$$(f*g)(t) = \frac{1}{\Gamma(n)}\int_0^t \tau^{n-1}e^{-\beta\tau}\mathrm{d}\tau.$$

If we make the change of variables $h = \beta\tau$ then

$$(f*g)(t) = \frac{1}{\Gamma(n)\beta^n}\int_0^{\beta t} h^{n-1}e^{-h}\mathrm{d}h. \equiv \mathrm{P}(n,\beta t),$$

which gives

$$\mathscr{L}^{-1}\left(\frac{\alpha_{(n)}^n}{s(\beta+s)^n}\right) = \left(\frac{\alpha_{(n)}}{\beta}\right)^n \mathrm{P}(n,\beta t). \tag{C.13}$$

To obtain the inverse of the second term of equation (C.12) we use again the properties of convolutions and the Laplace transform to get

$$\begin{aligned}\mathscr{L}^{-1}\left(\frac{\alpha_{(n)}^n\omega}{(s^2+\omega^2)(\beta+s)^n}\right) = &\frac{\alpha_{(n)}^n\sin(\omega t)}{\Gamma(n)}\int_0^t \tau^{n-1}e^{-\beta\tau}\cos(\omega\tau)\mathrm{d}\tau \\ &- \frac{\alpha_{(n)}^n\cos(\omega t)}{\Gamma(n)}\int_0^t \tau^{n-1}e^{-\beta\tau}\sin(\omega\tau)\mathrm{d}\tau.\end{aligned} \tag{C.14}$$

Solving integrals of (C.14) gives

$$\mathscr{L}^{-1}\left(\frac{\alpha_{(n)}^n\omega}{(s^2+\omega^2)(\beta+s)^n}\right) = \left(\frac{\alpha_{(n)}}{\omega^2+\beta^2}\right)^n \mathrm{Re}\left(\xi\varphi^n\overline{\mathrm{P}(n,\varphi t)}\right),$$

where $\varphi = \beta + i\omega$, $\xi = \sin(\omega t) + i\cos(\omega t)$, $\bar{z}$ is the complex conjugate of $z \in \mathbb{C}$, and $\mathrm{Re}(z)$ is the real part of z. It follows from the definition of the incomplete gamma function that $\overline{\mathrm{P}(n,\varphi t)} = \mathrm{P}(n,\bar{\varphi}t)$. We extract the real part of $\xi\varphi^n\overline{\mathrm{P}(n,\varphi t)}$ to get

$$\begin{aligned}&\mathscr{L}^{-1}\left(\frac{\alpha_{(n)}^n\omega}{(s^2+\omega^2)(\beta+s)^n}\right) = \\ &\left(\frac{\alpha_{(n)}}{r}\right)^n\left[\sin(\omega t+n\theta) - e^{-\beta t}\left\{\cos(n\theta)\sum_{k=0}^{n-1}\frac{t^k r^k\cos(k\theta)}{k!} - \sin(n\theta)\sum_{k=0}^{n-1}\frac{t^k r^k\sin(k\theta)}{k!}\right\}\right],\end{aligned}$$

where $r = (\beta^2 + \omega^2)^{1/2}$ and $\theta = \arctan\left(\frac{\beta}{\omega}\right)$. We now combine this last result with equation (C.13) the expression we have in equation (5.16):

$$\begin{aligned} x_n(t) &= \left(\frac{\alpha_{(n)}}{\beta}\right)^n \mathrm{P}(n,\beta t) + \left(\frac{\alpha_{(n)}}{r}\right)^n \Bigg[\sin(\omega t - n\theta) \\ &\quad - e^{-\beta t}\left\{ \cos(n\theta) \sum_{k=0}^{n-1} \frac{t^k r^k \cos(k\theta)}{k!} - \sin(n\theta) \sum_{k=0}^{n-1} \frac{t^k r^k \sin(k\theta)}{k!} \right\} \Bigg], \\ &= \left(\frac{\alpha_{(n)}}{\beta}\right)^n \mathrm{P}(n,\beta t) + \left(\frac{\alpha_{(n)}}{r}\right)^n \Bigg[\sin(\omega t - n\theta) \\ &\quad - e^{-\beta t}\left\{ \sum_{k=0}^{n} \frac{t^k r^k}{k!} \left(\cos(n\theta)\cos(k\theta) - \sin(n\theta)\sin(k\theta)\right) \right\} \Bigg], \\ &= \left(\frac{\alpha_{(n)}}{\beta}\right)^n \mathrm{P}(n,\beta t) + \left(\frac{\alpha_{(n)}}{r}\right)^n \left[\sin(\omega t - n\theta) - e^{-\beta t} \sum_{k=0}^{n} \frac{t^k r^k}{k!} \cos((n+k)\theta) \right] \end{aligned}$$

C.2 Properties of the ε-perturbed matrix of rates, $\mathbf{H}_i$

The matrix $\mathbf{H}_i$, which corresponds to a linear cascade with a perturbation ε at position i, is defined in Eq. (5.17) and has a Jordan decomposition given in Eq. (5.18):

$$\mathbf{H}_i = \mathbf{Q}_i \mathbf{J} \mathbf{Q}_i^{-1}. \tag{C.15}$$

As can be seen from the lower-triangular structure of $\mathbf{H}_i$, its Jordan form is the direct sum of two Jordan blocks associated with $-(\beta+\varepsilon)$, with multiplicity 1, and $-\beta$, with multiplicity $n-1$:

$$\mathbf{J} = \mathbf{J}_{-(\beta+\varepsilon)} \oplus \mathbf{J}_{-\beta}, \tag{C.16}$$

where

$$\mathbf{J}_{-(\beta+\varepsilon)} = \left[-(\beta+\varepsilon)\right]_{1\times 1}, \qquad \mathbf{J}_{-\beta} = \begin{bmatrix} -\beta & 1 & & \\ & -\beta & \ddots & \\ & & \ddots & 1 \\ & & & -\beta \end{bmatrix}_{(n-1)\times(n-1)}. \tag{C.17}$$

The transition matrix $\mathbf{Q}_i$ contains the generalised eigenvectors $\{\mathbf{q}_\ell^i\}_{\ell=1}^n$ as columns

$$\mathbf{Q}_i = \left[\mathbf{q}_1^i | \mathbf{q}_2^i | \dots | \mathbf{q}_n^i\right].$$

Proposition C.2.1. The following properties hold for $\mathbf{Q}_i$ and $\mathbf{J}$ (see Example C.2.1 for the case when $n = 6$):

I. $\mathbf{J}$ is independent of i.

II. $\mathbf{q}_1^i$, the eigenvector of $\mathbf{H}_i$ associated with $-(\beta + \varepsilon)$, is given by:

$$\mathbf{q}_1^i(j) = \begin{cases} 0 & \text{if } j < i, \\ \alpha_{(j)}^j \left(\frac{-1}{\varepsilon}\right)^{j-1} & \text{if } j \geq i. \end{cases} \tag{C.18}$$

III. $\{\mathbf{q}_{n-k+1}^i\}_{k=1}^{n-1}$, the $(n-1)$ generalised eigenvectors of $\mathbf{H}_i$ associated with $-\beta$, are given by:

- If $i \leq k$

$$\mathbf{q}_{n-k+1}^i(j) = \begin{cases} 0 & j \in \{1 \dots k\}, \\ -\alpha_{(j)}^j \left(\frac{-1}{\varepsilon}\right)^{j-k} & j \in \{k+1 \dots n\}. \end{cases} \tag{C.19}$$

- If $i > k$

$$\mathbf{q}_{n-k+1}^i(j) = \begin{cases} 0 & j \in \{1 \dots i-1\} \backslash k, \\ \alpha_{(k)}^k & j = k, \\ -\alpha_{(j)}^j \left(\frac{-1}{\varepsilon}\right)^{j-k} & j \in \{i \dots n\}. \end{cases} \tag{C.20}$$

IV. The inverse of the transition matrix $\mathbf{Q}_i^{-1}$ has the following structure:

- The first row is

$$\mathbf{Q}_i^{-1}(1,j) = \begin{cases} \frac{(-\varepsilon)^{j-1}}{\alpha_{(j)}^j} & \text{if } j \leq i, \\ 0 & \text{if } j > i. \end{cases} \tag{C.21}$$

- Rows $k \in \{2, \dots, n\}$ are

$$\mathbf{Q}_i^{-1}(k,j) = \begin{cases} 0 & \text{if } j \leq n-k, \\ \frac{1}{\alpha_{(j)}^j} & \text{if } j = n-k+1, \\ \frac{\varepsilon}{\alpha_{(j)}^j} & \text{if } j = n-k+2 \text{ and } n-i+1 \geq k, \\ 0 & \text{otherwise.} \end{cases} \tag{C.22}$$

V. For all i:

$$\mathbf{Q}_i^{-1}\mathbf{e}_1 = \mathbf{e}_1 + \mathbf{e}_n = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 1 \end{bmatrix}. \tag{C.23}$$

Proof. I. All the matrices $\mathbf{H}_i$ have the same eigenvalues with the same multiplicity so the Jordan form $\mathbf{J}$ is identical for all i.

II. Let $\mathbf{v}_1^i = \mathbf{H}_i\mathbf{q}_1^i$, if $i = 1$ it is straightforward to see that $\mathbf{v}_1^1(1) = -(\beta+\varepsilon)$ and when $j > i$

$$\mathbf{v}_1^1(j) = \alpha_j\alpha_{(j-1)}^{j-1}\left(\frac{-1}{\varepsilon}\right)^{j-2} - \beta\alpha_{(j)}^j\left(\frac{-1}{\varepsilon}\right)^{j-1} = -(\beta+\varepsilon)\alpha_{(j)}^j\left(\frac{-1}{\varepsilon}\right)^{j-1}, \tag{C.24}$$

so $\mathbf{v}_1^1 = -(\beta+\varepsilon)\mathbf{q}_1^1$.
When $i > 1$ then $\mathbf{v}_1^i(j) = 0$ for $j = 1 \dots i-1$. If $j = i$, then

$$\mathbf{v}_1^j(i) = -(\beta+\varepsilon)\alpha_{(i)}^i\left(\frac{-1}{\varepsilon}\right)^{i-1},$$

and if $j > i$ then the same situation as in equation (C.24) applies, so again we have $\mathbf{v}_1^i = -(\beta+\varepsilon)\mathbf{q}_1^i$.

III. Define $\mathbf{B}_i = (\mathbf{H}_i + \beta\mathbf{I}_n)$, we can see from the definition of $\mathbf{q}_2^i$ in equations (C.19) and (C.20) that $\mathbf{B}_i\mathbf{q}_2^i = \mathbf{0}$, ie $\mathbf{q}_2^i$ is the eigenvector of $\mathbf{H}_i$ associated to $-\beta$. The rest of the generalised eigenvectors $\mathbf{q}_3^i \dots \mathbf{q}_n^i$ associated with $-\beta$ can be multiplied by $\mathbf{B}_i$

to show that

$$\mathbf{B}_i \mathbf{q}_h^i = \mathbf{q}_{h-1}^i, \qquad h = 3 \dots n. \tag{C.25}$$

It follows that $\mathbf{B}_i^{h-1} \mathbf{q}_h^i = \mathbf{0}$, so $\mathbf{Q}_i$ as defined in equations (C.18), (C.19), and (C.20) is the transition matrix of generalised eigenvectors of $\mathbf{H}_i$.

IV. We can verify that $\mathbf{Q}_i^{-1}$ defined in equations in equations (C.21) and (C.22) is the inverse of the transition matrix $\mathbf{Q}_i$ by multiplying them to obtain $\mathbf{Q}_i^{-1}\mathbf{Q}_i = \mathbf{I}_n$.

V. This property follows directly from the structure of $\mathbf{Q}_i^{-1}$.

□

Example C.2.1. Consider a cascade of length $n = 6$. The structures of the transition matrix and its inverse for $i = 1, \dots, 6$ are

$$\mathbf{Q}_1 = \begin{bmatrix} \alpha_{(1)} & 0 & 0 & 0 & 0 & 0 \\ -\frac{\alpha_{(2)}^2}{\varepsilon} & 0 & 0 & 0 & 0 & \frac{\alpha_{(2)}^2}{\varepsilon} \\ \frac{\alpha_{(3)}^3}{\varepsilon^2} & 0 & 0 & 0 & \frac{\alpha_{(3)}^3}{\varepsilon} & -\frac{\alpha_{(3)}^3}{\varepsilon^2} \\ -\frac{\alpha_{(4)}^4}{\varepsilon^3} & 0 & 0 & \frac{\alpha_{(4)}^4}{\varepsilon} & -\frac{\alpha_{(4)}^4}{\varepsilon^2} & \frac{\alpha_{(4)}^4}{\varepsilon^3} \\ \frac{\alpha_{(5)}^5}{\varepsilon^4} & 0 & \frac{\alpha_{(5)}^5}{\varepsilon} & -\frac{\alpha_{(5)}^5}{\varepsilon^2} & \frac{\alpha_{(5)}^5}{\varepsilon^3} & -\frac{\alpha_{(5)}^5}{\varepsilon^4} \\ -\frac{\alpha_{(6)}^6}{\varepsilon^5} & \frac{\alpha_{(6)}^6}{\varepsilon} & -\frac{\alpha_{(6)}^6}{\varepsilon^2} & \frac{\alpha_{(6)}^6}{\varepsilon^3} & -\frac{\alpha_{(6)}^6}{\varepsilon^4} & \frac{\alpha_{(6)}^6}{\varepsilon^5} \end{bmatrix}, \quad \mathbf{Q}_1^{-1} = \begin{bmatrix} \frac{1}{\alpha_{(1)}} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{\alpha_{(5)}^5} & \frac{\varepsilon}{\alpha_{(6)}^6} \\ 0 & 0 & 0 & \frac{1}{\alpha_{(4)}^4} & \frac{\varepsilon}{\alpha_{(5)}^5} & 0 \\ 0 & 0 & \frac{1}{\alpha_{(3)}^3} & \frac{\varepsilon}{\alpha_{(4)}^4} & 0 & 0 \\ 0 & \frac{1}{\alpha_{(2)}^2} & \frac{\varepsilon}{\alpha_{(3)}^3} & 0 & 0 & 0 \\ \frac{1}{\alpha_{(1)}} & \frac{\varepsilon}{\alpha_{(2)}^2} & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{C.26}$$

$$\mathbf{Q}_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \alpha_{(1)} \\ -\frac{\alpha_{(2)}^2}{\varepsilon} & 0 & 0 & 0 & 0 & \frac{\alpha_{(2)}^2}{\varepsilon} \\ \frac{\alpha_{(3)}^3}{\varepsilon^2} & 0 & 0 & 0 & \frac{\alpha_{(3)}^3}{\varepsilon} & -\frac{\alpha_{(3)}^3}{\varepsilon^2} \\ -\frac{\alpha_{(4)}^4}{\varepsilon^3} & 0 & 0 & \frac{\alpha_{(4)}^4}{\varepsilon} & -\frac{\alpha_{(4)}^4}{\varepsilon^2} & \frac{\alpha_{(4)}^4}{\varepsilon^3} \\ \frac{\alpha_{(5)}^5}{\varepsilon^4} & 0 & \frac{\alpha_{(5)}^5}{\varepsilon} & -\frac{\alpha_{(5)}^5}{\varepsilon^2} & \frac{\alpha_{(5)}^5}{\varepsilon^3} & -\frac{\alpha_{(5)}^5}{\varepsilon^4} \\ -\frac{\alpha_{(6)}^6}{\varepsilon^5} & \frac{\alpha_{(6)}^6}{\varepsilon} & -\frac{\alpha_{(6)}^6}{\varepsilon^2} & \frac{\alpha_{(6)}^6}{\varepsilon^3} & -\frac{\alpha_{(6)}^6}{\varepsilon^4} & \frac{\alpha_{(6)}^6}{\varepsilon^5} \end{bmatrix}, \quad \mathbf{Q}_2^{-1} = \begin{bmatrix} \frac{1}{\alpha_{(1)}} & -\frac{\varepsilon}{\alpha_{(2)}^2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{\alpha_{(5)}^5} & \frac{\varepsilon}{\alpha_{(6)}^6} \\ 0 & 0 & 0 & \frac{1}{\alpha_{(4)}^4} & \frac{\varepsilon}{\alpha_{(5)}^5} & 0 \\ 0 & 0 & \frac{1}{\alpha_{(3)}^3} & \frac{\varepsilon}{\alpha_{(4)}^4} & 0 & 0 \\ 0 & \frac{1}{\alpha_{(2)}^2} & \frac{\varepsilon}{\alpha_{(3)}^3} & 0 & 0 & 0 \\ \frac{1}{\alpha_{(1)}} & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{C.27}$$

$$\mathbf{Q}_3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \alpha_{(1)} \\ 0 & 0 & 0 & 0 & \alpha & 0 \\ \frac{\alpha^3_{(3)}}{\varepsilon^2} & 0 & 0 & 0 & \frac{\alpha^3_{(3)}}{\varepsilon} & -\frac{\alpha^3_{(3)}}{\varepsilon^2} \\ -\frac{\alpha^4_{(4)}}{\varepsilon^3} & 0 & 0 & \frac{\alpha^4_{(4)}}{\varepsilon} & -\frac{\alpha^4_{(4)}}{\varepsilon^2} & \frac{\alpha^4_{(4)}}{\varepsilon^3} \\ \frac{\alpha^5_{(5)}}{\varepsilon^4} & 0 & \frac{\alpha^5_{(5)}}{\varepsilon} & -\frac{\alpha^5_{(5)}}{\varepsilon^2} & \frac{\alpha^5_{(5)}}{\varepsilon^3} & -\frac{\alpha^5_{(5)}}{\varepsilon^4} \\ -\frac{\alpha^6_{(6)}}{\varepsilon^5} & \frac{\alpha^6_{(6)}}{\varepsilon} & -\frac{\alpha^6_{(6)}}{\varepsilon^2} & \frac{\alpha^6_{(6)}}{\varepsilon^3} & -\frac{\alpha^6_{(6)}}{\varepsilon^4} & \frac{\alpha^6_{(6)}}{\varepsilon^5} \end{bmatrix}, \quad \mathbf{Q}_3^{-1} = \begin{bmatrix} \frac{1}{\alpha_{(1)}} & -\frac{\varepsilon}{\alpha^2_{(2)}} & \frac{\varepsilon^2}{\alpha^3_{(3)}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{\alpha^5_{(5)}} & \frac{\varepsilon}{\alpha^6_{(6)}} \\ 0 & 0 & 0 & \frac{1}{\alpha^4_{(4)}} & \frac{\varepsilon}{\alpha^5_{(5)}} & 0 \\ 0 & 0 & \frac{1}{\alpha^3_{(3)}} & \frac{\varepsilon}{\alpha^4_{(4)}} & 0 & 0 \\ 0 & \frac{1}{\alpha^2_{(2)}} & 0 & 0 & 0 & 0 \\ \frac{1}{\alpha_{(1)}} & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{C.28}$$

$$\mathbf{Q}_4 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \alpha_{(1)} \\ 0 & 0 & 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & \alpha^3_{(3)} & 0 & 0 \\ -\frac{\alpha^4_{(4)}}{\varepsilon^3} & 0 & 0 & \frac{\alpha^4_{(4)}}{\varepsilon} & -\frac{\alpha^4_{(4)}}{\varepsilon^2} & \frac{\alpha^4_{(4)}}{\varepsilon^3} \\ \frac{\alpha^5_{(5)}}{\varepsilon^4} & 0 & \frac{\alpha^5_{(5)}}{\varepsilon} & -\frac{\alpha^5_{(5)}}{\varepsilon^2} & \frac{\alpha^5_{(5)}}{\varepsilon^3} & -\frac{\alpha^5_{(5)}}{\varepsilon^4} \\ -\frac{\alpha^6_{(6)}}{\varepsilon^5} & \frac{\alpha^6_{(6)}}{\varepsilon} & -\frac{\alpha^6_{(6)}}{\varepsilon^2} & \frac{\alpha^6_{(6)}}{\varepsilon^3} & -\frac{\alpha^6_{(6)}}{\varepsilon^4} & \frac{\alpha^6_{(6)}}{\varepsilon^5} \end{bmatrix}, \quad \mathbf{Q}_4^{-1} = \begin{bmatrix} \frac{1}{\alpha_{(1)}} & -\frac{\varepsilon}{\alpha^2_{(2)}} & \frac{\varepsilon^2}{\alpha^3_{(3)}} & -\frac{\varepsilon^3}{\alpha^4_{(4)}} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{\alpha^5_{(5)}} & \frac{\varepsilon}{\alpha^6_{(6)}} \\ 0 & 0 & 0 & \frac{1}{\alpha^4_{(4)}} & \frac{\varepsilon}{\alpha^5_{(5)}} & 0 \\ 0 & 0 & \frac{1}{\alpha^3_{(3)}} & 0 & 0 & 0 \\ 0 & \frac{1}{\alpha^2_{(2)}} & 0 & 0 & 0 & 0 \\ \frac{1}{\alpha_{(1)}} & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{C.29}$$

$$\mathbf{Q}_5 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \alpha_{(1)} \\ 0 & 0 & 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & \alpha^3_{(3)} & 0 & 0 \\ 0 & 0 & \alpha^4_{(4)} & 0 & 0 & 0 \\ \frac{\alpha^5_{(5)}}{\varepsilon^4} & 0 & \frac{\alpha^5_{(5)}}{\varepsilon} & -\frac{\alpha^5_{(5)}}{\varepsilon^2} & \frac{\alpha^5_{(5)}}{\varepsilon^3} & -\frac{\alpha^5_{(5)}}{\varepsilon^4} \\ -\frac{\alpha^6_{(6)}}{\varepsilon^5} & \frac{\alpha^6_{(6)}}{\varepsilon} & -\frac{\alpha^6_{(6)}}{\varepsilon^2} & \frac{\alpha^6_{(6)}}{\varepsilon^3} & -\frac{\alpha^6_{(6)}}{\varepsilon^4} & \frac{\alpha^6_{(6)}}{\varepsilon^5} \end{bmatrix}, \quad \mathbf{Q}_5^{-1} = \begin{bmatrix} \frac{1}{\alpha_{(1)}} & -\frac{\varepsilon}{\alpha^2_{(2)}} & \frac{\varepsilon^2}{\alpha^3_{(3)}} & -\frac{\varepsilon^3}{\alpha^4_{(4)}} & \frac{\varepsilon^4}{\alpha^5_{(5)}} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{\alpha^5_{(5)}} & \frac{\varepsilon}{\alpha^6_{(6)}} \\ 0 & 0 & 0 & \frac{1}{\alpha^4_{(4)}} & 0 & 0 \\ 0 & 0 & \frac{1}{\alpha^3_{(3)}} & 0 & 0 & 0 \\ 0 & \frac{1}{\alpha^2_{(2)}} & 0 & 0 & 0 & 0 \\ \frac{1}{\alpha_{(1)}} & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{C.30}$$

$$\mathbf{Q}_6 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \alpha_{(1)} \\ 0 & 0 & 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & \alpha^3_{(3)} & 0 & 0 \\ 0 & 0 & \alpha^4_{(4)} & 0 & 0 & 0 \\ 0 & \alpha^5_{(5)} & 0 & 0 & 0 & 0 \\ -\frac{\alpha^6_{(6)}}{\varepsilon^5} & \frac{\alpha^6_{(6)}}{\varepsilon} & -\frac{\alpha^6_{(6)}}{\varepsilon^2} & \frac{\alpha^6_{(6)}}{\varepsilon^3} & -\frac{\alpha^6_{(6)}}{\varepsilon^4} & \frac{\alpha^6_{(6)}}{\varepsilon^5} \end{bmatrix}, \quad \mathbf{Q}_6^{-1} = \begin{bmatrix} \frac{1}{\alpha_{(1)}} & -\frac{\varepsilon}{\alpha^2_{(2)}} & \frac{\varepsilon^2}{\alpha^3_{(3)}} & -\frac{\varepsilon^3}{\alpha^4_{(4)}} & \frac{\varepsilon^4}{\alpha^5_{(5)}} & -\frac{\varepsilon^5}{\alpha^6_{(6)}} \\ 0 & 0 & 0 & 0 & \frac{1}{\alpha^5_{(5)}} & 0 \\ 0 & 0 & 0 & \frac{1}{\alpha^4_{(4)}} & 0 & 0 \\ 0 & 0 & \frac{1}{\alpha^3_{(3)}} & 0 & 0 & 0 \\ 0 & \frac{1}{\alpha^2_{(2)}} & 0 & 0 & 0 & 0 \\ \frac{1}{\alpha_{(1)}} & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{C.31}$$

C.3 Calculation of $x_{n+1}(t)$ with one ε-perturbed inactivation rate

C.3.1 Constant stimulus, derivation of Eq. (5.25)

We solve the differential equation (5.24) through an integrating factor to get:

$$e^{(\beta+\varepsilon)t}x_{n+1}(t) = \alpha_{n+1}\left(\frac{\alpha_{(n)}}{\beta}\right)^n \int e^{(\beta+\varepsilon)t}\mathrm{P}(n,\beta t)\mathrm{d}t + c. \tag{C.32}$$

Use the properties of the gamma function to re-express the integral as

$$\int e^{(\beta+\varepsilon)t}\mathrm{P}(n,\beta t)\mathrm{d}t = \int e^{(\beta+\varepsilon)t}\mathrm{d}t - \int e^{\varepsilon t}\sum_{k=0}^{n-1}\frac{(\beta t)^k}{k!}\mathrm{d}t. \tag{C.33}$$

and solve the second integral of (C.33) using integration by parts:

$$\int e^{\varepsilon t}\sum_{k=0}^{n-1}\frac{(\beta t)^k}{k!}\mathrm{d}t = \frac{e^{\varepsilon t}}{\varepsilon}\sum_{k=0}^{n-1}\frac{(\beta t)^k}{k!} - \int \frac{\beta e^{\varepsilon t}}{\varepsilon}\left(\sum_{k=0}^{n-1}\frac{(\beta t)^k}{k!} - \frac{(\beta t)^{n-1}}{(n-1)!}\right)\mathrm{d}t \tag{C.34}$$

to obtain

$$\left(1+\frac{\beta}{\varepsilon}\right)\int e^{\varepsilon t}\sum_{k=0}^{n-1}\frac{(\beta t)^k}{k!}\mathrm{d}t = \frac{e^{\varepsilon t}}{\varepsilon}\sum_{k=0}^{n-1}\frac{(\beta t)^k}{k!} + \frac{\beta^n}{\varepsilon(n-1)!}\int e^{\varepsilon t}t^{n-1}\mathrm{d}t. \tag{C.35}$$

The integral on the right-hand side of equation (C.35) can be solved using the formula (Gradshteyn and Ryzhik, 2007):

$$\int e^{\varepsilon t}t^{n-1}\mathrm{d}t = e^{\varepsilon t}\left(\sum_{k=0}^{n-1}\frac{(-1)^k(n-1)!}{\varepsilon^{k+1}(n-1-k)!}t^{n-1-k}\right). \tag{C.36}$$

Substituting in equation (C.33) and gathering terms we obtain

$$\int e^{(\beta+\varepsilon)t}\mathrm{P}(n,\beta t)\mathrm{d}t = \frac{e^{\varepsilon t}}{\beta+\varepsilon}\left(e^{\beta t} - \sum_{k=0}^{n-1}\frac{(\beta t)^k}{k!} + \frac{(-1)^k\beta^n t^{n-1-k}}{\varepsilon^{k+1}(n-1-k)!}\right) + c, \tag{C.37}$$

whence Eq. (C.32) becomes

$$x_{n+1}(t) = \frac{\alpha_{n+1}}{\beta+\varepsilon}\left(\frac{\alpha_{(n)}}{\beta}\right)^n \left(1 - e^{-\beta t}\sum_{k=0}^{n-1}\left[\frac{(\beta t)^k}{k!} + \frac{(-1)^k \beta^n t^{n-1-k}}{\varepsilon^{k+1}(n-1-k)!}\right]\right) + ce^{-(\beta+\varepsilon)t}. \quad \text{(C.38)}$$

The initial condition $x_{n+1}(0) = 0$ requires that

$$c = (-1)^{n+1}\frac{\alpha_{n+1}}{\beta+\varepsilon}\left(\frac{\alpha_{(n)}}{\varepsilon}\right)^n, \quad \text{(C.39)}$$

which gives the final expression given in equation (5.25):

$$x_{n+1}(t) = \frac{\alpha_{n+1}}{\beta+\varepsilon}\left(\frac{\alpha_{(n)}}{\beta}\right)^n \left(1 - e^{-\beta t}\left[\left(\frac{-\beta}{\varepsilon}\right)^n e^{-\varepsilon t} + \sum_{k=0}^{n-1}\frac{(\varepsilon^{n-k} - (-\beta)^{n-k})(\beta t)^k}{\varepsilon^{n-k}k!}\right]\right). \quad \text{(C.40)}$$

C.3.2 Exponentially decaying stimulus, derivation of Eqs. (5.29) and (5.30)

Consider first the case $\beta \neq \lambda$. To solve the differential equation (5.26), define $\sigma = \beta - \lambda$ and use integrating factors to get:

$$e^{(\beta+\varepsilon)t}x_{n+1} = \alpha_{n+1}\left(\frac{\alpha_{(n)}}{\sigma}\right)^n \left(\int e^{(\sigma+\varepsilon)t}\mathrm{d}t - \int e^{-\varepsilon t}\sum_{k=0}^{n-1}\frac{(\sigma t)^k}{k!}\mathrm{d}t\right) + c. \quad \text{(C.41)}$$

Using integration by parts for the second integral on the right-hand side, and following similar steps to those above gives

$$e^{(\beta+\varepsilon)t}x_{n+1} = \frac{\alpha_{n+1}}{\sigma+\varepsilon}\left(\frac{\alpha_{(n)}}{\sigma}\right)^n \left(e^{(\sigma+\varepsilon)t} - e^{\varepsilon t}\sum_{k=0}^{n-1}\frac{(\varepsilon^{n-k} - (-\sigma)^{n-k})(\sigma t)^k}{\varepsilon^{n-k}k!}\right) + c. \quad \text{(C.42)}$$

The initial condition $x_{n+1}(0) = 0$ requires that

$$c = \frac{\alpha_{n+1}}{\sigma+\varepsilon}\left(\frac{\alpha_{(n)}}{\varepsilon}\right)^n, \quad \text{(C.43)}$$

thus giving equation (5.29):

$$x_{n+1}(t) = \frac{\alpha_{n+1}}{\beta+\varepsilon-\lambda}\left(\frac{\alpha_{(n)}}{\beta-\lambda}\right)^n \left[e^{-\lambda t} + \frac{e^{-(\beta+\varepsilon)t}}{\varepsilon^n} - \right.$$
$$\left. -e^{-\beta t}\sum_{k=0}^{n-1}\frac{\left(\varepsilon^{n-k}-(\lambda-\beta)^{n-k}\right)(\beta-\lambda)^k t^k}{\varepsilon^{n-k}k!}\right]. \qquad \text{(C.44)}$$

When $\lambda = \beta$ we use integrating factors to solve equation (5.27):

$$x_{n+1}(t)e^{(\beta+\varepsilon)t} = \frac{\alpha_{(n+1)}}{\Gamma(n+1)}\int t^n e^{\varepsilon t}\mathrm{d}t + c \;\; = \frac{\alpha_{(n+1)}}{\Gamma(n+1)}\frac{n!}{\varepsilon}\sum_{k=0}^{n}\frac{(-1)^k t^{(n-k)}}{\varepsilon^k(n-k)!} + c, \quad \text{(C.45)}$$

and taking the initial condition $x_{n+1}(0) = 0$ we get equation (5.30):

$$x_{n+1}(t) = \left(\frac{\alpha_{(n+1)}}{\varepsilon}\right)^{n+1} e^{-\beta t}\left[\varepsilon^n\sum_{k=0}^{n}\frac{(-1)^k t^{n-k}}{\varepsilon^k(n-k)!} + (-1)^{n+1}e^{-\varepsilon t}\right].$$

C.4 Parameter fitting

We use the optimisation method developed in Chapter 4 to fit the parameters of our models to the artificial 'observed data' used in Examples 5.4.1 and 5.4.3.

C.4.1 Model simplification and parameter fitting

In Example 5.4.1 we used Algorithm 4.1 from Chapter 4 to obtain the parameter values of the gamma function expressions (5.12) and (5.14) from data taken from numerical solutions of equation (5.3) with both constant and decaying inputs and random added noise (see Fig. 5.5).

In the example we considered two cascades with parameters $n = 5$, $\alpha_1 = 3$, $\alpha_i = 4$ $i = 2, \ldots, 5$, and $\beta = 3$ from resting initial conditions. The first of the cascades has input $R(t) = \alpha_1$, the second; $R(t) = \alpha_1 e^{-\lambda t}$, $\lambda = 1$. To avoid confusion, in this section we denote the output of the cascade with constant stimulus as $x_{5c}(t)$, and the output from using a decaying stimulus as $x_{5d}(t)$. Table C.1 shows both the untouched sample from the numerical solutions of the equations of both cascades (dashed lines in both

t	$x_{5c}(t)$	$\tilde{x}_{5c}(t)$	$x_{5d}(t)$	$\tilde{x}_{5d}(t)$
0	0	0	0	0
1	0.584	0.5821	0.4651	0.4667
2	2.2611	2.2977	1.205	1.2423
3	2.9855	3.1372	0.8544	0.9596
4	3.1339	3.1167	0.3963	0.3352
5	3.1573	3.2542	0.1572	0.1451
6	3.1601	3.332	0.059	0.068
7	3.1607	3.0908	0.0218	0.0244
8	3.1606	3.1167	0.008	0.0071
9	3.1604	3.2505	0.003	0.0034
10	3.1604	3.2208	0.0011	0.0009

Table C.1: Data sampled from numerically solving equation (5.3) for constant and decaying inputs (columns $x_{5c}(t)$ and $x_{5d}(t)$) using parameters $n = 5$, $\alpha_1 = 3$, $\alpha_i = 4$, $\beta = 3$ and $\lambda = 1$, and by adding random noise to the solutions (columns $\tilde{x}_{5c}(t)$ and $\tilde{x}_{5d}(t)$).

plots of Fig. 5.5B), and the sample with added random noise sampled from a $\mathcal{N}(0, 0.05^2)$ distribution (squares in Fig. 5.5B), which we denote by $\tilde{x}_{5c}$ and $\tilde{x}_{5d}$.

We define the parameter vectors $\boldsymbol{\theta}_c = [\alpha_{(n)},\ \beta,\ n]$ for the cascade with constant stimulus, and $\boldsymbol{\theta}_d = [\alpha_{(n)},\ \beta,\ n,\ \lambda]$ for the cascade with decaying stimulus. To find the estimates of $\boldsymbol{\theta}_c$ and $\boldsymbol{\theta}_d$ we minimise the error (4.3) on each cascade:

$$E_{\mathcal{D}}(\boldsymbol{\theta}_c) = \sum_{t=0}^{10} ||x_{5c}(t; \boldsymbol{\theta}_c) - \tilde{x}_{5c}(t)||_2^2 \,, \tag{C.46}$$

$$E_{\mathcal{D}}(\boldsymbol{\theta}_d) = \sum_{t=0}^{10} ||x_{5d}(t; \boldsymbol{\theta}_d) - \tilde{x}_{5d}(t)||_2^2 \,, \tag{C.47}$$

where $x_{5c}(t; \boldsymbol{\theta}_c)$, $x_{5d}(t; \boldsymbol{\theta}_d)$ are functions (5.12) and (5.14) evaluated at $\boldsymbol{\theta}_c$ and $\boldsymbol{\theta}_d$, respectively.

We begin each minimisation with a prior distribution $\mathcal{U}(0, 10)$ for each parameter. At every iteration of the algorithm we simulate 500 points in the parameter space, to which we apply local minimisation. Of the 500 local minima we select the 50 with the smallest error to construct the posterior. We sample 500 new points from the posterior (now a prior) and continue the sample-minimise-cull cycle until the means of the errors of the posteriors in

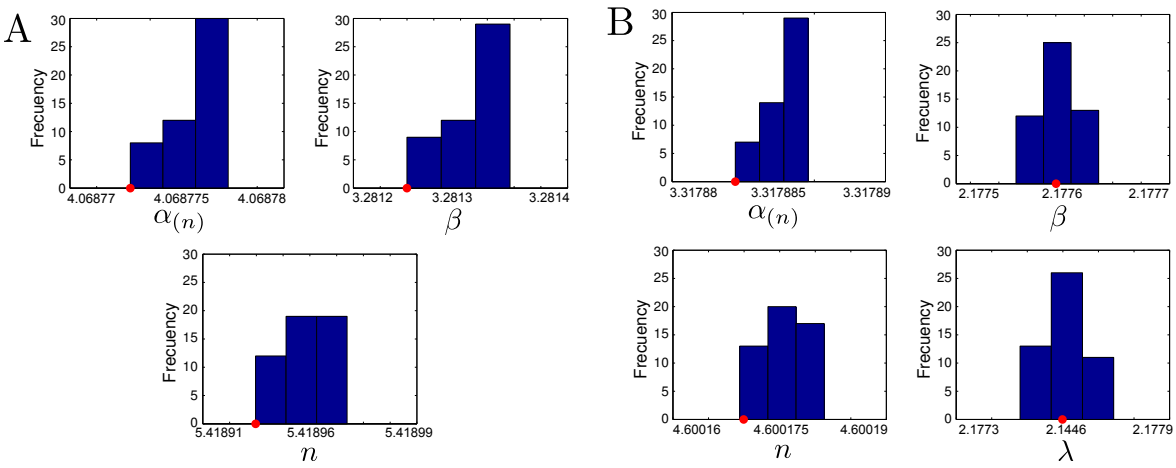

Figure C.1: **A**: Histograms of the parameters of equation (5.12) found by minimising error (C.46). **B**: Histograms of the fit of equation (5.14) to error (C.47). In all histograms the red dot marks the value of the parameter that gave the smallest error (ie the components of $\boldsymbol{\theta}_c^\ddagger$ and $\boldsymbol{\theta}_d^\ddagger$).

consecutive iterations (ϕ_k) is smaller that 10^{-5} (see Chapter 4 for details on the method). Figure C.1 shows the histograms of the components of $\boldsymbol{\theta}_c^\dagger$ and $\boldsymbol{\theta}_d^\dagger$ obtained after minimising errors (C.46) and (C.47). On each case the algorithm converged after only two iterations, with $\phi_2 \approx 0$. The values of the parameters of the cascade with constant stimulus are $\alpha_{(n)} \approx 4.068$, $\beta \approx 3.281$, and $n \approx 5.418$, giving $E_{\mathcal{D}}(\boldsymbol{\theta}_c^\ddagger) = 0.2338$. The parameters obtained with the decaying stimulus are $\alpha_{(n)} \approx 3.317$, $\beta \approx 2.177$, $n \approx 4.6$, and $\lambda \approx 2.177$, with error $E_{\mathcal{D}}(\boldsymbol{\theta}_d^\ddagger) = 0.0842$.

C.4.2 Delay differential equation models for activation cascades

In Example 5.4.3 we approximate data drawn from the solution of a delay differential equation with a linear cascade represented by an incomplete gamma function expression. We use Algorithm 4.1 to find the parameter values that minimise the distance from the data to equation (5.12).

We simulate the DDE system (5.31) with parameters $\hat{\alpha} = 2$, $\hat{\beta} = 3$, and $\tau = 2$ from resting initial conditions (see Fig. 5.7B, top). We denote the solution to equation (5.31) by $\hat{p}_2(t)$ and the solution with added random noise from a $\mathcal{N}(0, 0, 05^2)$ distribution by $\tilde{p}_2(t)$, shown in Table C.2. With equation (5.32) we approximate $\hat{p}_2(t)$ using Algorithm 4.1 to

t	$\hat{p}_2(t)$	$\tilde{p}_2(t)$
0	0	0
0.2524	0	0.0091
0.5336	0	0.1521
0.9385	0	0.0019
1.2083	0	0.1227
1.55	0	0.0348
2.0069	0.0002	0.001
2.5374	0.6565	0.5782
2.8395	1.1261	1.1848
3.0717	1.4221	1.397
3.4184	1.7484	1.7964
3.7092	1.9289	1.9957
4	2.0488	2.041
4.9736	2.2142	2.3031
5.4868	2.2362	2.4671
6	2.2447	2.2972
7	2.251	2.2498
8	2.2496	2.3409
9	2.2501	2.2557
10	2.25	2.1393

Table C.2: Data sampled from a numerical solution of equation (5.31) ($\hat{p}_2(t)$) using parameters $\hat{\alpha} = 2$, $\hat{\beta} = 3$, and $\tau = 2$. Column $\tilde{p}_2(t)$ contains the numerical solution data with added random noise.

find the parameters $\boldsymbol{\theta} = [\alpha_{(n)},\ \beta,\ n]$ that minimise the error

$$E_{\mathcal{D}}(\boldsymbol{\theta}) = \sum_{i=1}^{20} ||p_n(t_i; \boldsymbol{\theta}) - \tilde{p}_2(t_i)||_2^2 .$$

We use the prior $\mathcal{U}(0, 10)$ for all the parameters. As before, we sample 500 points from the parameter space, minimise them locally and use the 50 best local minima to construct a posterior. We continue the sample-minimise-cull interactions until $\phi_k \leq 10^{-5}$. The algorithm finished after two iterations ($\phi_2 \approx 0$), obtaining parameters $\alpha_{(n)} \approx 2.27$, $\beta \approx 7.53$, and $n \approx 22.1072$. The posterior distribution of all parameters is extremely narrow (support centred on each parameter value $\pm 10^{-8}$).

C.5 Cascades with negative feedback

C.5.1 Constant stimulus

Consider the cascade given by the system of ODEs:

$$\begin{aligned} \frac{\mathrm{d}x_1}{\mathrm{d}t} &= R(t) - \beta x_1 - \epsilon x_1 x_n, \\ \frac{\mathrm{d}x_i}{\mathrm{d}t} &= \alpha_i x_{i-1} - \beta x_i, \qquad i = 2, \ldots, n, \end{aligned} \tag{C.48}$$

where $x_1(0) = x_i(0) = 0$. When $R(t) = \alpha_1$ we can try to approximate the solutions by expanding x_k in powers of ϵ:

$$x_k = x_{k0} + \epsilon x_{k1} + \epsilon^2 x_{k2} + \ldots .$$

From equation (5.12) we get the leading-order approximation to x_k as a normalised lower-incomplete gamma function:

$$x_{k0}(t) = \left(\frac{\alpha_{(k)}}{\beta}\right)^k \mathrm{P}(k, \beta t).$$

When $k = 1$ we get the order 1 correction by solving the equation:

$$\begin{aligned}\frac{\mathrm{d}x_{11}}{\mathrm{d}t} &= -\beta x_{11} - x_{10}x_{n0},\\ &= -\beta x_{11} - \frac{\alpha_1 \alpha_{(n)}^n}{\beta^{n+1}}\mathrm{P}(1,\beta t)\mathrm{P}(b,\beta t).\end{aligned}$$

The order 1 approximation to $x_1(t)$ is then

$$x_1(t) \approx \frac{\alpha_1}{\beta}\mathrm{P}(1,\beta t) - \frac{\epsilon\alpha_1\alpha_{(n)}^n}{\beta^{n+2}}\left[\mathrm{P}(n,\beta t) + \left\{n\mathrm{P}(n+1,\beta t) - \frac{(\beta t)^n}{n!} - \beta t\mathrm{P}(n,\beta t)\right\}\right]e^{-\beta t}.$$

C.5.2 Exponentially decreasing stimulus

When $R(t) = \alpha_1 e^{-\lambda t}$ in equation (C.48) we can again expand $x_n(t)$ in powers of ϵ and obtain the leading order approximation in terms of the incomplete gamma function (5.14):

$$x_{n0}(t) = \begin{cases} \left(\frac{\alpha_{(n)}}{\beta-\lambda}\right)^n e^{-\lambda t}\,\mathrm{P}(n,(\beta-\lambda)t) & \text{if } \beta \neq \lambda \\ \\ \frac{1}{\Gamma(n+1)}\left(\alpha_{(n)}t\right)^n e^{-\beta t} & \text{if } \beta = \lambda, \end{cases}$$

To obtain $x_{n1}(t)$ lets suppose that $\beta = \lambda$; furthermore, we assume that a suitable nondimensionalisation has been performed so that

$$x_{n0}(t) = \frac{t^n e^{-t}}{n!}.$$

The order-one correction can be shown to be:

$$x_{n1}(t) = (-1)^{n+1}e^{-2t}(t+n) - e^{-t}\sum_{k=1}^{n}(-1)^n\frac{kt^{n-k}}{(n-k)!}.$$

The sum in this last equation is a particular case of the hypergeometric function:

$$\sum_{k=1}^{n}(-1)^n\frac{kt^{n-k}}{(n-k)!} = \frac{t^{n-1}}{(n-1)!}\sum_{k=1}^{n}\left\{\frac{(-1)^{k+1}k}{t^{k-1}}\prod_{j=1}^{k}(n-j)\right\},$$
$$= \frac{t^{n-1}}{(n-1)!}\sum_{m=0}^{n-1}\left\{\frac{(-1)^{m}(m+1)}{t^{m}}\prod_{j=1}^{m-1}(n-j)\right\},$$

changing variables $m = k-1$, and $(-1)^{k+1} = (-1)^{k-1} = k^m$

$$\ldots = \frac{t^{n-1}}{(n-1)!}\sum_{m=0}^{n-1}\left\{\frac{m+1}{t^{m}}\prod_{j=1}^{m+1}(-m+j)\right\},$$
$$= \frac{t^{n-1}}{(n-1)!}\sum_{m=0}^{n-1}\left\{\frac{(m+1)\Gamma(-n+1+m)}{t^{m}\Gamma(-n+1)}\right\},$$

using $\Gamma(z+1) = z\Gamma(z)$ recursively, also $\Gamma(2) = 1$ so:

$$\ldots = \frac{t^{n-1}}{(n-1)!}\sum_{m=0}^{n-1}\left\{\frac{m+1}{t^{m}}\frac{m!}{m!}\frac{\Gamma(-n+1+m)}{\Gamma(2)\Gamma(-n+1)}\right\},$$
$$= \frac{t^{n-1}}{(n-1)!}\sum_{m=0}^{n-1}\left\{\frac{\Gamma(-n+1+m)\Gamma(2+m)}{\Gamma(-n+1)\Gamma(2)}\frac{t^{-m}}{m!}\right\}. \tag{C.49}$$

Once that $-n+1+m \geq 2$, we cannot cancel out $\Gamma(-n+1)$ out of the denominator of equation (C.49). Because $\Gamma(-n+1)$ goes to infinity when $-n+1$ is a negative integer, all terms containing it in the denominator go to zero, so

$$\ldots = \frac{t^{n-1}}{(n-1)!}\sum_{m=0}^{\infty}\left\{\frac{\Gamma(-n+1+m)\Gamma(2+m)}{\Gamma(-n+1)\Gamma(2)}\frac{t^{-m}}{m!}\right\},$$
$$= \frac{t^{n-1}}{(n-1)!}{}_2F_0\left([-n+1,2];,t^{-1}\right),$$

Where ${}_2F_0$ is a case of the hypergeometric defined as (Abramowitz and Stegun, 1964):

$${}_2F_0([a,b];\ ,z) = \sum_{j=0}^{\infty}(a)_j(b)_j\frac{z^j}{j!},$$

and $(a)_j$ is Pochhammer's symbol:

$$(a)_j = \frac{\Gamma(a+j)}{\Gamma(a)}.$$

Therefore the order-one approximation to $X_n(t)$ is

$$x_n(t) \approx \frac{t^n}{n!}e^{-t} + \varepsilon_n \left[(-1)^{n+1}e^{-2t}(t+n) - \frac{e^{-t}t^{n-1}}{(n-1)!}{}_2F_0\left([-n+1,2];,t^{-1}\right)\right]. \quad \text{(C.50)}$$

The results described above can be used to try to approximate the behaviour of activation cascades in a signalling network, eg when only the output of the cascade can be measured.

C.6 A result from Golub and Van Loan (1996)

In Golub and Van Loan (1996) the following statement is provided as an exercise[1]:
Show that if

$$\exp\left(z\begin{bmatrix} -\mathbf{A}^T & \mathbf{P} \\ \mathbf{0} & \mathbf{A} \end{bmatrix}\right) = \begin{bmatrix} \mathbf{F}_{11} & \mathbf{F}_{12} \\ \mathbf{0} & \mathbf{F}_{22} \end{bmatrix},$$

where $\mathbf{A}, \mathbf{P} \in \mathbb{R}^{n\times n}$ and $z \in \mathbb{R}$, then

$$\mathbf{F}_{22}^T\mathbf{F}_{12} = \int_0^z e^{s\mathbf{A}^T}\mathbf{P}e^{s\mathbf{A}}\mathrm{d}s.$$

A proof is given by Van Loan (1976), here we show an alternative version using the definition of the matrix exponential that arose from studying approximations to nonlinear activation cascades of length n.
It can be easily verified from the definition of $e^{\mathbf{A}}$) that $\mathbf{F}_{22} = e^{z\mathbf{A}}$ and

$$\mathbf{F}_{12} = \sum_{k=1}^{\infty}\sum_{i=0}^{k-1}\frac{z^k}{k!}\left(-\mathbf{A}^T\right)^i\mathbf{P}\left(\mathbf{A}\right)^{k-1-i},$$

[1]Page 577 problem 11.34. In the book it is incorrectly printed as $F_{11}^TF_{12} = \int_0^z \ldots$, it should be $F_{22}^TF_{12}$.

so we have

$$\mathbf{F}_{22}^{T}\mathbf{F}_{12} = e^{z\mathbf{A}^{T}} \sum_{k=1}^{\infty}\sum_{i=0}^{k-1} \frac{z^{k}}{k!} \left(-\mathbf{A}^{T}\right)^{i} \mathbf{P} \left(\mathbf{A}\right)^{k-1-i}.$$

Differentiating the last expression gives

$$\frac{\mathrm{d}\mathbf{F}_{22}^{T}\mathbf{F}_{12}}{\mathrm{d}z} = \mathbf{A}^{T} e^{z\mathbf{A}^{T}} \sum_{k=1}^{\infty}\sum_{i=0}^{k-1} \frac{z^{k}}{k!} \left(-\mathbf{A}^{T}\right)^{i} \mathbf{P} \left(\mathbf{A}\right)^{k-1-i} + e^{z\mathbf{A}^{T}} \sum_{k=1}^{\infty}\sum_{i=0}^{k-1} \frac{z^{k-1}}{(k-1)!} \left(-\mathbf{A}^{T}\right)^{i} \mathbf{P} \left(\mathbf{A}\right)^{k-1-i},$$

rearranging indices and using the fact that a matrix and its exponential commute, then

$$= e^{z\mathbf{A}^{T}} \left[\sum_{k=0}^{\infty}\sum_{i=0}^{k} \frac{z^{k}}{k!} \left(-\mathbf{A}^{T}\right)^{i} \mathbf{P} \left(\mathbf{A}\right)^{k-i} - \sum_{k=1}^{\infty}\sum_{i=1}^{k} \frac{z^{k}}{k!} \left(-\mathbf{A}^{T}\right)^{i} \mathbf{P} \left(\mathbf{A}\right)^{k-i} \right],$$

where only the terms with $i = 0$ or $k = 0$ do not get cancelled out and

$$\begin{aligned} \frac{\mathrm{d}\mathbf{F}_{22}^{T}\mathbf{F}_{12}}{\mathrm{d}z} &= e^{zA}\mathbf{P} \left[\mathbf{I}_{n} + z\mathbf{A} + \frac{z^{2}}{2}\mathbf{A} + \ldots \right] \\ &= e^{z\mathbf{A}^{T}}\mathbf{P}e^{z\mathbf{A}}, \end{aligned}$$

which is the result we wanted to obtain.

Appendix D

Additional modelling and computations

D.1 Henderson-Hasselbach equation

Let the concentrations of protons H^+, a weak acid A^-, and protonated acid HA in a solute be $[H^+]$, $[A^-]$, and $[HA]$, respectively. The reaction

$$H^+ + A^- \underset{k_2}{\overset{k_1}{\rightleftharpoons}} HA,$$

is represented by the ODE

$$\frac{\mathrm{d}[H^+]}{\mathrm{d}t} = k_2[HA] - k_1[H^+][A^-].$$

The equilibrium concentration of H^+ is $[H^+] = K_a \frac{[HA]}{[A]}$, where $K_a = \frac{k_2}{k_1}$. If we take a logarithm (base 10) on both sides and multiply by -1 we get

$$-\log([H^+]) = -\log(K_a) + \log\left(\frac{[A^-]}{[HA]}\right). \tag{D.1}$$

The right-hand side of equation (D.1) is equivalent to pH according to the Henderson-Hasselbach equation (Po and Senozan, 2001); this means that

$$pH = -\log([H^+]).$$

D.2 Proton transport across the vacuolar membrane

In simple terms the transport across the vacuolar membrane via the V–ATPase can be seen as the reaction (Keener and Sneyd, 1998):

$$H^+_{cyt} + ATP \underset{k_2}{\overset{k_1}{\rightleftharpoons}} H^+_{vac} + ADP + P.$$

If $[H^+]_{cyt}$ and $[H^+]_{vac}$ are the proton concentrations in the cytoplasm and the vacuole, the ODE model describing the change in proton concentration in the cytoplasm is

$$\frac{\mathrm{d}[H^+]_{cyt}}{\mathrm{d}t} = -k_1[H^+]_{cyt}[ATP] + k_2[H^+]_{vac}[ADP][P], \tag{D.2}$$

and the Gibbs free-energy for the reaction is:

$$G_\triangle = RT\ln\left(\frac{[H^+]_{vac}[ADP][P]}{[H^+]_{cyt}[ATP]}\right) - FE_{m_{vac}} + G_{\triangle ATP},$$

where R is the gas constant, T; the absolute temperature (in K), F; the Faraday constant, $E_{m_{vac}}$; the vacuolar membrane potential (given by the Nernst equation), and $G_{\triangle ATP}$; the Gibbs free-energy for ATP-hydrolysis (Ke, 2010). When $G_\triangle = 0$ (energy equilibrium) we have

$$\frac{[H^+]_{vac}[ADP][P]}{[H^+]_{cyt}[ATP]} = \exp\left(-\frac{G_{\triangle ATP} - FE_{m_{vac}}}{RT}\right).$$

From equation (D.2) at steady state we obtain the relation

$$k_1[H^+]_{cyt}[ATP] = k_2[H^+]_{vac}[ADP][P],$$

which gives

$$\frac{k_1}{k_2} = \exp\left(-\frac{G_{\triangle ATP} - FE_{m_{vac}}}{RT}\right).$$

The flux through the channel is described by (Ke, 2010; Keener and Sneyd, 1998):

$$J_{V\text{-}ATPase} = k_1\left\{[H^+]_{cyt}[ATP] - [H^+]_{vac}[ADP][P]\exp\left(-\frac{G_{\triangle ATP} - FE_{m_{vac}}}{RT}\right)\right\}.$$

In Sec. 2.3.4 we discussed that V–ATPase activity is increased by active OST1, a result

of ABA-inactivation of ABI1, which means that the rate of proton transport from the cytoplasm to the vacuoles has a dependency on ABA. We can argue that the dependency on ABA (because it is mediated by enzymes) can saturate and thus be approximated with a Michaelis-Menten term, for example:

$$k_1([ABA]) = \hat{k}_1 + \frac{\tilde{k}_1[ABA]}{\kappa + [ABA]}.$$

D.3 Fitting the parameters of the stomatal closure model

Equations (7.22)-(7.27) have 28 parameters whose values must be determined. (Note that the variables are rescaled dividing them by 100, so that control levels and initial conditions are 1, to improve numerical stability of the fitting process.)

We reduce the number of free parameters from 28 to 23 through a series of assumptions. Firstly, we assume that in the absence of stimuli (ie $[ABA] = [ETH] = 0$) we must remain stationary in the control state:

$$\frac{\mathrm{d}[ROS]}{\mathrm{d}t} = \frac{\mathrm{d}[AOX]}{\mathrm{d}t} = \frac{\mathrm{d}[NO]}{\mathrm{d}t} = \frac{\mathrm{d}[K^+_{out}]}{\mathrm{d}t} = \frac{\mathrm{d}[K^+]}{\mathrm{d}t} = 0, \tag{D.3}$$

Hence the following relationships between parameters must hold:

$$\begin{aligned} \alpha_{10} &= \beta_{11}, \\ \alpha_{20} &= \beta_{20}, \\ \beta_{30} &= \frac{\alpha_{31}}{k_{31}+1}, \\ \beta_{40} &= \alpha_{40} + \alpha_{42}, \\ \beta_{50} &= \frac{\alpha_{51}}{k_{51}+1}. \end{aligned} \tag{D.4}$$

With these conditions, the number of unknown parameters has been reduced to 23 and we

define the vector $\boldsymbol{\theta} \in \mathbb{R}^{23}$ in parameter space:

$$\begin{aligned}\boldsymbol{\theta} = [&\alpha_{11}, \alpha_{12}, k_{11}, k_{12}, \beta_{11}, \beta_{12}, \beta_{13}, n_2, \alpha_{22}, \alpha_{23}, k_{21}, k_{22},\\ &\alpha_{24}, n_1, \beta_{20}, \alpha_{31}, k_{31}, \alpha_{32}, \alpha_{41}, \alpha_{42}, \alpha_{43}, \alpha_{51}, k_{51}]\,,\end{aligned}$$

whose components are all non-negative. We define the treatments as the following set: $\mathcal{T} = \{T_1, T_2, T_3, T_4, T_5, T_6, T_7\}$, where

$$\begin{aligned}T_1 &= 1\,\mu M \text{ ABA} + 0\,\mu M \text{ ethylene},\\ T_2 &= 0\,\mu M \text{ ABA} + 1\,\mu M \text{ ethylene},\\ T_3 &= 10\,\mu M \text{ ABA} + 0\,\mu M \text{ ethylene},\\ T_4 &= 0\,\mu M \text{ ABA} + 10\,\mu M \text{ ethylene},\\ T_5 &= 10\,\mu M \text{ ABA} + 10\,\mu M \text{ ethylene}.\\ T_6 &= 50\,\mu M \text{ ABA} + 0\,\mu M \text{ ethylene}.\\ T_7 &= 0\,\mu M \text{ ABA} + 100\,\mu M \text{ ethylene}.\end{aligned}$$

The data set $\mathcal{D}$ with measurements of ROS and aperture consists of the observations shown in Figs. 6.1 and 6.2. We denote by $[\widehat{ROS}]_i$ and $[\widehat{AP}]_i$ the vectors of measurements of ROS and aperture under treatment T_i; and by $[ROS]_i(\boldsymbol{\theta})$ and $[AP]_i(\boldsymbol{\theta})$ the model predictions at the same time points as the data with dose T_i, using parameters $\boldsymbol{\theta}$. The discrepancy of the model in equations (7.22)-(7.27) and the data is measured by the following error function:

$$E_{\mathcal{D}}(\boldsymbol{\theta}) = \sum_{i\in\{1,2,3\}} \left|\left|[\widehat{ROS}]_i - [ROS]_i(\boldsymbol{\theta})\right|\right|_2^2 + \left|\left|[\widehat{AP}]_i - [AP]_i(\boldsymbol{\theta})\right|\right|_2^2, \tag{D.5}$$

where $||\cdot||_2$ is the euclidean norm. That is, we measure the distance between our ROS and aperture measurements and the model for a given $\boldsymbol{\theta}$ in the parameter space. The global optimisation problem is to find $\boldsymbol{\theta}^{\ddagger}$ where

$$\begin{aligned}\boldsymbol{\theta}^{**} &= \min_{\boldsymbol{\theta}} E_{\mathcal{D}}(\boldsymbol{\theta}),\\ &\text{subject to } \boldsymbol{\theta} \geq 0.\end{aligned}$$

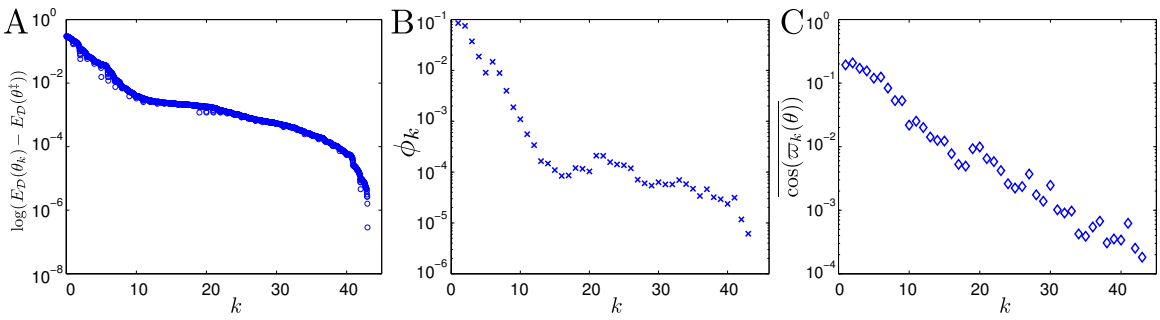

Figure D.1: Convergence criteria for optimisation algorithm. **A**: Difference between the best 50 local minima from each iteration (k) of the algorithm (the posterior) and the global minimum $\boldsymbol{\theta}^{\ddagger}$ on a semilogarithmic scale. **B**: Difference between the mean of the errors of the local minima from consecutive iterations of the fitting algorithm (ϕ_k) on a semilogarithmic scale. **C**: Mean of the cosines among individual local minima from each iteration on a semilogarithmic scale.

We use the Squeeze-and-breathe optimisation method from Chapter 4 to find $\boldsymbol{\theta}^{\ddagger}$. The method requires an initial probability distribution for each parameter (called a *prior*). Here we use a uniform distribution $U(0, 10)$ for all parameters. On each iteration 500 points in the parameter space (in $\mathbb{R}_{+}^{23}$) are sampled from the prior. Each point is used as a starting point to minimise $E_{\mathcal{D}}(\boldsymbol{\theta})$ locally (using the Nelder-Mead simplex algorithm). The 50 local minima with the smallest errors are used to construct a *posterior* distribution of the parameters. The posterior is used as a prior for the next iteration where another 500 points are sampled and minimised until the convergence criteria has been met. Figure D.1 shows the convergence of the method for fitting the parameters of our model. On Fig. D.1A, we show the decrease in the difference (on a semilogarithmic scale) between the errors of the parameter sets found at the end of each iteration and the global minimum $E_{\mathcal{D}}(\boldsymbol{\theta}^{\ddagger}) \approx 0.0215$, obtained at the end of iteration 43. The 50 lowest errors of each iteration minus $E_{\mathcal{D}}(\boldsymbol{\theta}^{\ddagger})$ are shown on a decreasing order from left to right. Figure D.1B shows the convergence criterion defined in Chapter 4. We stop the iterations of the method once the difference between mean of the errors of the 50 parameter sets from consecutive iterations (ϕ_k, shown on a semilogarithmic scale) is smaller that 10^{-5}. During the first 20 iterations of the method ϕ_k appears to decrease exponentially and thereafter the trend still continues downwards albeit no longer exponentially. On Fig. D.1C we show the mean of the cosines of the

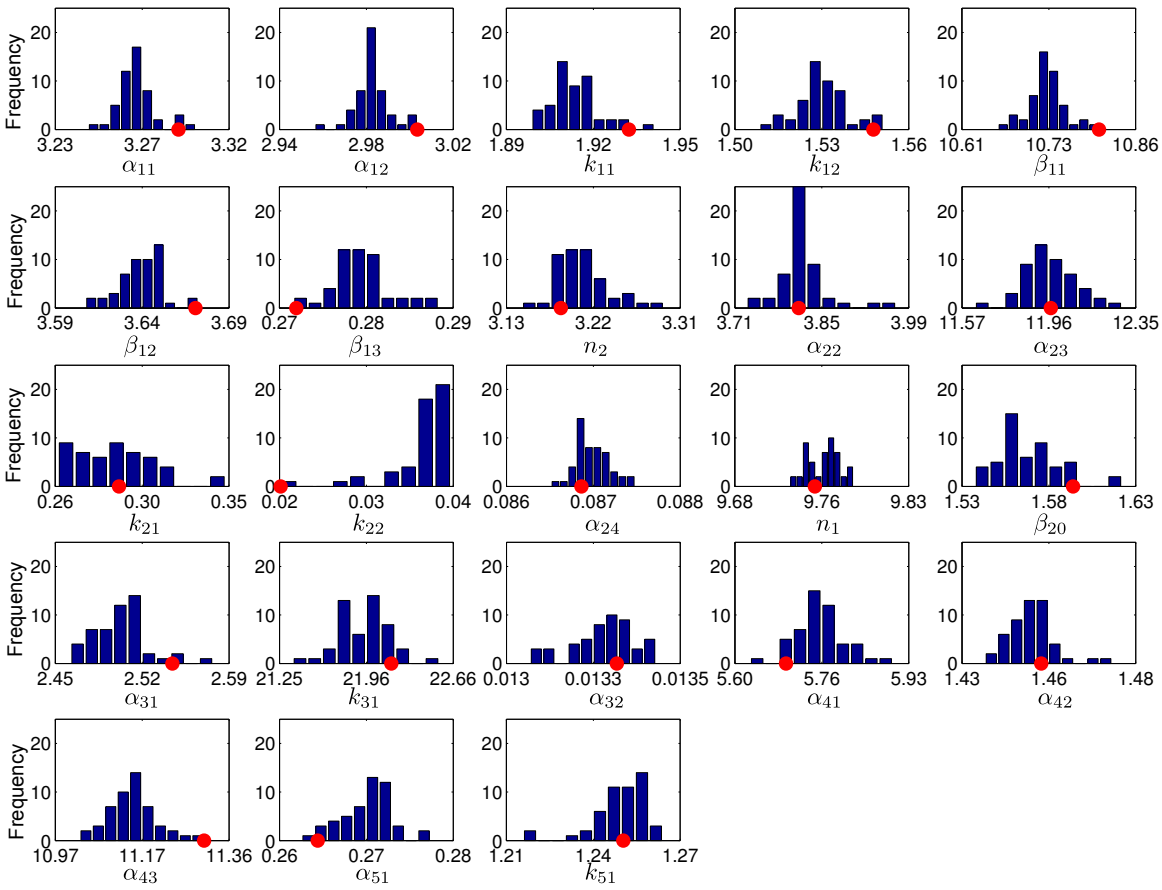

Figure D.2: Frequencies of the parameters after fitting. Histograms of the distribution of the 50 best parameters obtained after 43 iterations of the fitting algorithm. The red dots indicate the value of the parameter that gave the smallest error ($\boldsymbol{\theta}^{\ddagger}$).

angles between all local minima from each iteration. This is to assert that the method converges to a single region of the parameter space. After iteration 43 the mean cosine is $O(10^{-4})$. Based on these metrics we conclude that $\boldsymbol{\theta}^{\ddagger}$ is a good estimation of the model parameters, given the present data. Figure D.2 shows the distribution of the best 50 parameters after 43 iterations of the algorithm. Red dots mark the mean of each parameter (values in Table D.1). The behaviour of the model that we observe in Figs. 7.7–7.10 is given by these parameters.

Parameter	Value	Units
α_{11}	3.2976	min^{-1}
α_{12}	3.0054	min^{-1}
k_{11}	1.9389	dimensionless
k_{12}	1.551	dimensionless
β_{11}	10.8119	min^{-1}
β_{12}	3.6752	min^{-1}
β_{13}	0.2749	dimensionless
n_2	3.192	dimensionless
α_{21}	3.8179	min^{-1}
α_{22}	11.9759	dimensionless
k_{21}	0.2962	dimensionless
k_{22}	0.0268	dimensionless
α_{23}	0.0868	dimensionless
n_1	9.755	dimensionless
β_{20}	1.5973	min^{-1}
α_{31}	2.5528	min^{-1}
k_{31}	22.1635	dimensionless
α_{32}	0.0133	min^{-1}
α_{40}	5.7029	min^{-1}
α_{41}	1.4584	min^{-1}
α_{42}	11.3133	min^{-1}
α_{51}	0.2713	min^{-1}
k_{51}	1.256	dimensionless
α_{10}	10.8119	min^{-1}
α_{20}	1.5973	min^{-1}
β_{30}	0.1102	min^{-1}
β_{40}	17.0162	min^{-1}
β_{50}	0.1203	min^{-1}

Table D.1: Parameter values obtained by the Squeeze-and-Breathe algorithm. Other values are given by the relations given in the text. The bottom five parameters are calculated using the expressions in (D.4).

References

Abramoff, M. D., Magelhaes, P. J., and Ram, S. J. (2004). Image Processing with ImageJ. *Biophotonics International*, **11**(7), 36–42.

Abramowitz, M. and Stegun, I. A. (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York, 9 edition.

Acharya, B. and Assmann, S. (2009). Hormone interactions in stomatal function. *Plant Molecular Biology*, **69**, 451–462. 10.1007/s11103-008-9427-0.

Aldana, M. (2003). Boolean dynamics of networks with scale-free topology. *Physica D: Nonlinear Phenomena*, **185**(1), 45 – 66.

Allen, G. J., Kuchitsu, K., Chu, S. P., Murata, Y., and Schroeder, J. I. (1999). Arabidopsis abi1-1 and abi2-1 phosphatase mutations reduce abscisic acid-induced cytoplasmic calcium rises in guard cells. *Plant Cell*, **11**(9), 1785–1798.

Alon, U. (2007). *An introduction to systems biology: design principles of biological circuits*. Chapman and Hall/CRC mathematical & computational biology series. Chapman & Hall/CRC.

Anderson, R. M. and May, R. M. (1992). *Infectious Diseases of Humans Dynamics and Control*. Oxford University Press.

Apel, K. and Hirt, H. (2004). Reactive Oxygen Species: Metabolism, Oxidative Stress, and Signal Transduction. *Annual Review of Plant Biology*, **55**(1), 373–399.

Assmann and Shimazaki (1999). The multisensory guard cell. Stomatal responses to blue light and abscisic acid . *Plant Physiology*, **119**(3), 809–816.

August, E., Parker, K. H., and Barahona, M. (2007). A dynamical model of lipoprotein metabolism. *Bull Math Biol*, **69**(4), 1233–1254.

Bai, X., Yang, L., Tian, M., Chen, J., Shi, J., Yang, Y., and Hu, X. (2011). Nitric Oxide Enhances Desiccation Tolerance of Recalcitrant *Antiaris toxicaria* Seeds via Protein S-Nitrosylation and Carbonylation. *PLoS ONE*, **6**(6), e20714.

Bar-Or, R. L., Maya, R., Segel, L. A., Alon, U., Levine, A. J., and Oren, M. (2000). Generation of oscillations by the p53-Mdm2 feedback loop: a theoretical and experimental study. *Proc Natl Acad Sci USA*, **97**(21), 11250–11255.

Bari, R. and Jones, J. D. G. (2009). Role of plant hormones in plant defence responses. *Plant Molecular Biology*, **69**(4), 473–488.

Belin, C., de Franco, P.-O., Bourbousse, C., Chaignepain, S., Schmitter, J.-M., Vavasseur, A., Giraudat, J., Barbier-Brygoo, H., and Thomine, S. (2006). Identification of Features Regulating OST1 Kinase Activity and OST1 Function in Guard Cells. *Plant Physiology*, **141**(4), 1316–1327.

Bellman, R., Bellman, R., and Cooke, K. (1963). *Differential-difference equations*. Mathematics in science and engineering. Academic Press.

Bender, C. M. and Orszag, S. A. (1999). *Advanced Mathematical Methods for Scientists and Engineers. Asymptotic Methods and Perturbation Theory*. Springer.

Bernard, S., Cajavec, B., Pujo-Menjouet, L., Mackey, M. C., and Herzel, H. (2006). Modelling transcriptional feedback loops: the role of Gro/TLE1 in Hes1 oscillations. *Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences*, **364**(1842), 1155–1170.

Bethke, P. C., Libourel, I. G. L., and Jones, R. L. (2006). Nitric oxide reduces seed dormancy in Arabidopsis. *J Exp Bot*, **57**(3), 517–526.

Blatt, M. R. and Armstrong, F. (1993). K+ channels of stomatal guard cells: Abscisic-acid-evoked control of the outward rectifier mediated by cytoplasmic pH. *Planta*, **191**(3), 330–341.

Bliss, R. D., Painter, P. R., and Marr, A. G. (1982). Role of feedback inhibition in stabilizing the classical operon. *Journal of Theoretical Biology*, **97**(2), 177 – 193.

Brewer, D., Barenco, M., Callard, R., Hubank, M., and Stark, J. (2008). Fitting ordinary differential equations to short time course data. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **366**(1865), 519–544.

Brewer, D. S. (2006). *Modelling the p53 Gene Regulatory Network*. Ph.D. thesis, University of London.

Bright, J. (2006). *Nitric oxide signalling in Arabidopsis thaliana guard cells*. Ph.D. thesis, University of the West of England.

Bright, J., Desikan, R., Hancock, J. T., Weir, I. S., and Neill, S. J. (2006). ABA-induced NO generation and stomatal closure in Arabidopsis are dependent on H_2O_2 synthesis. *The Plant Journal*, **45**(1), 113–122.

Brown, K. S. and Sethna, J. P. (2003). Statistical mechanical approaches to models with many poorly known parameters. *Physical Review E*, **68**(2), 021904.

Carter, G. W., Rush, C. G., Uygun, F., Sakhanenko, N. A., Galas, D. J., and Galitski, T. (2010). A systems-biology approach to modular genetic complexity. *Chaos*, **20**(2), 026102.

Chan, C. C., Stark, J., and George, A. J. (1999). Analysis of cytokine dynamics in corneal allograft rejection. *Proc Biol Sci*, **266**(1434), 2217–2223.

Chang, C. and Bleecker, A. B. (2004). Ethylene Biology. More Than a Gas. *Plant Physiology*, **136**(2), 2895–2899.

Chang, L. and Karin, M. (2001). Mammalian MAP kinase signalling cascades. *Nature*, **410**(6824), 37–40.

Chaves, M., Sontag, E. D., and Dinerstein, R. J. (2004). Optimal Length and Signal Amplification in Weakly Activated Signal Transduction Cascades. *The Journal of Physical Chemistry B*, **108**(39), 15311–15320.

Chen, W. W., Niepel, M., and Sorger, P. K. (2010). Classic and contemporary approaches to modeling biochemical reactions. *Genes Dev*, **24**(17), 1861–1875.

Chen, Y.-F., Etheridge, N., and Schaller, G. E. (2005). Ethylene Signal Transduction. *Ann Bot*, **95**(6), 901–915.

Chen, Y.-L., Huang, R., Xiao, Y.-M., Lu, P., Chen, J., and Wang, X.-C. (2004). Extracellular Calmodulin-Induced Stomatal Closure Is Mediated by Heterotrimeric G Protein and H2O2. *Plant Physiology*, **136**(4), 4096–4103.

Chen, Z. and Gallie, D. R. (2004). The Ascorbic Acid Redox State Controls Guard Cell Signaling and Stomatal Movement. *Plant Cell*, **16**(5), 1143–1162.

Cho, D., Kim, S. A., Murata, Y., Lee, S., Jae, S.-K., Nam, H. G., and Kwak, J. M. (2009a). De-regulated expression of the plant glutamate receptor homolog AtGLR3.1 impairs long-term Ca2+-programmed stomatal closure. *The Plant Journal*, **58**(3), 437–449.

Cho, D., Shin, D., Jeon, B., and Kwak, J. (2009b). ROS-Mediated ABA Signaling. *Journal of Plant Biology*, **52**(2), 102–113.

Colijn, C. and Mackey, M. C. (2005). A mathematical model of hematopoiesis–I. Periodic chronic myelogenous leukemia. *Journal of Theoretical Biology*, **237**(2), 117–132.

Conzelmann, H., Saez-Rodriguez, J., Sauter, T., Bullinger, E., Allgower, F., and Gilles, E. (2004). Reduction of mathematical models of signal transduction networks: simulation-based approach applied to EGF receptor signalling. *Systems Biology, IEE Proceedings*, **1**(1), 159–169.

Cornish-Bowden, A. (2004). *Fundamentals of enzyme kinetics*. Portland Press.

Csikász-Nagy, A., Cardelli, L., and Soyer, O. S. (2010). Response dynamics of phosphorelays suggest their potential utility in cell signalling. *Journal of The Royal Society Interface*.

D'Angelo, C., Weinl, S., Batistic, O., Pandey, G. K., Cheong, Y. H., Schltke, S., Albrecht, V., Ehlert, B., Schulz, B., Harter, K., Luan, S., Bock, R., and Kudla, J. (2006). Alternative complex formation of the ca2+-regulated protein kinase cipk1 controls abscisic acid-dependent and independent stress responses in arabidopsis. *The Plant Journal*, **48**(6), 857–872.

Delledonne, M., Zeier, J., Marocco, A., and Lamb, C. (2001). Signal interactions between nitric oxide and reactive oxygen intermediates in the plant hypersensitive disease resistance response. *Proc Natl Acad Sci U S A*, **98**(23), 13454–13459.

Desikan, R., Griffiths, R., Hancock, J., and Neill, S. (2002). A new role for an old enzyme: nitrate reductase-mediated nitric oxide generation is required for abscisic acid-induced stomatal closure in Arabidopsis thaliana. *Proc Natl Acad Sci USA*, **99**(25), 16314–16318.

Desikan, R., Hancock, J. T., Bright, J., Harrison, J., Weir, I., Hooley, R., and Neill, S. J. (2005). A Role for ETR1 in Hydrogen Peroxide Signaling in Stomatal Guard Cells. *Plant Physiology*, **137**(3), 831–834.

Desikan, R., Last, K., Harrett-Williams, R., Tagliavia, C., Harter, K., Hooley, R., Hancock, J. T., and Neill, S. J. (2006). Ethylene-induced stomatal closure in Arabidopsis occurs via AtrbohF-mediated hydrogen peroxide synthesis. *The Plant Journal*, **47**(6), 907–916.

Desikan, R., Hancock, J., and Neill, S. (2007). *Reactive Oxygen Species as Signalling Molecules*, pages 169–196. Blackwell Publishing Ltd.

Devienne, E. (2010). Nitric Oxide is a Novel Component of Ethylene Cell Signalling in Arabidopsis Guard Cells. BSc in Biology, Final Year Project, Imperial College London.

Díaz, J. and Álvarez-Buylla, E. (2006). A model of the ethylene signaling pathway and its gene response in Arabidopsis thaliana: Pathway cross-talk and noise-filtering properties. *Chaos*, **16**(2), 023112.

Díaz, J. and Álvarez-Buylla, E. (2009). Information flow during gene activation by signaling molecules: ethylene transduction in Arabidopsis cells as a study system. *BMC Systems Biology*, **3**(1), 48.

Dittrich, P. and Raschke, K. (1977). Malate metabolism in isolated epidermis of Commelina communis L. in relation to stomatal functioning. *Planta*, **134**(1), 77–81.

Edelstein-Keshet, L. (1988). *Mathematical models in biology*. Classics in applied mathematics. SIAM.

Elowitz, M. B. and Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. *Nature*, **403**(6767), 335–338.

Feliu, E., Andersen, L. N., Knudsen, M., and Wiuf, C. (2010). A General Mathematical Framework Suitable for Studying Signaling Cascades. *arXiv:1008.0427*.

Feliú, E., Knudsen, M., Andersen, L., and Wiuf, C. (2011). An Algebraic Approach to Signaling Cascades with n Layers. *Bulletin of Mathematical Biology*, pages 1–28. 10.1007/s11538-011-9658-0.

Finkel, T. and Holbrook, N. J. (2000). Oxidants, oxidative stress and the biology of ageing. *Nature*, **408**(6809), 239–247.

Fowler, A. (1997). *Mathematical models in the applied sciences*. Cambridge texts in applied mathematics. Cambridge University Press.

Freeman, S. (2008). *Biological Science*. Pearson, 3 edition.

Fujioka, A., Terai, K., Itoh, R. E., Aoki, K., Nakamura, T., Kuroda, S., Nishida, E., and Matsuda, M. (2006). Dynamics of the ras/erk mapk cascade as monitored by fluorescent probes. *J Biol Chem*, **281**(13), 8917–8926.

García-Mata, C. and Lamattina, L. (2002). Nitric oxide and abscisic acid cross talk in guard cells. *Plant Physiology*, **128**(3), 790–792.

García-Mata, C. and Lamattina, L. (2007). Abscisic acid (ABA) inhibits light-induced stomatal opening through calcium- and nitric oxide-mediated signaling pathways. *Nitric Oxide*, **17**(3-4), 143 – 151.

García-Mata, C., Gay, R., Sokolovski, S., Hills, A., Lamattina, L., and Blatt, M. R. (2003). Nitric oxide regulates K+ and Cl- channels in guard cells through a subset of abscisic acid-evoked signaling pathways. *Proc Natl Acad Sci USA*, **100**(19), 11116–11121.

Gedney, N., Cox, P. M., Betts, R. A., Boucher, O., Huntingford, C., and Stott, P. A. (2006). Detection of a direct carbon dioxide effect in continental river runoff records. *Nature*, **439**(7078), 835–838.

Geiger, D., Scherzer, S., Mumm, P., Stange, A., Marten, I., Bauer, H., Ache, P., Matschi, S., Liese, A., Al-Rasheid, K. A. S., Romeis, T., and Hedrich, R. (2009). Activity of guard cell anion channel SLAC1 is controlled by drought-stress signaling kinase-phosphatase pair. *Proc Natl Acad Sci USA*, **106**(50), 21425–21430.

Gershenfeld, N. (1999). *The nature of mathematical modeling*. Cambridge University Press.

Gershenson, C. (2003). Classification of random boolean networks. In *Proceedings of the eighth international conference on Artificial Life*, pages 1–8, Cambridge, MA, USA. MIT Press.

Gershenson, C. (2004). Introduction to Random Boolean Networks. *arXiv:0408006*.

Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, **81**(25), 2340–2361.

Golub, G. H. and Van Loan, C. F. (1996). *Matrix Computations*. Johns Hopkins Studies in Mathematical Sciences. The Johns Hopkins University Press, 3rd edition.

Gosti, F., Beaudoin, N., Serizet, C., Webb, A. A. R., Vartanian, N., and Giraudat, J. (1999). ABI1 Protein Phosphatase 2C Is a Negative Regulator of Abscisic Acid Signaling. *Plant Cell*, **11**(10), 1897–1910.

Gradshteyn, I. and Ryzhik, I. (2007). *Table of Integrals, Series, and Products*. Academic Press.

Gross, J. L. and Yellen, J., editors (2004). *Handbook of Graph Theory*. CRC Press.

Gunawardena, J. (2011). A linear elimination framework. *arXiv:1109.6231*.

Guo, H. and Ecker, J. R. (2004). The ethylene signaling pathway: new insights. *Current Opinion in Plant Biology*, **7**(1), 40 – 49.

Gutenkunst, R. N., Waterfall, J. J., Casey, F. P., Brown, K. S., Myers, C. R., and Sethna, J. P. (2007). Universally Sloppy Parameter Sensitivities in Systems Biology Models. *PLoS Comput Biol*, **3**(10), e189.

Hancock, J. T. (1999). *Cell Signalling*. Longman, 1 edition.

Hardy, K. and Stark, J. (2002). Mathematical models of the balance between apoptosis and proliferation. *Apoptosis*, **7**(4), 373–381.

Harrington, H. A., Komorowski, M., Beguerisse Díaz, M., Ratto, G. M., and Stumpf, M. P. H. (2011). Dependence of MAPK mediated signaling on Erk isoforms and differences in nuclear shuttling. *arXiv:1110.3742*. Submitted to Physical Biology.

Hedrich, R., Busch, H., and Raschke, K. (1990). Ca2+ and nucleotide dependent regulation of voltage dependent anion channels in the plasma membrane of guard cells. *EMBO J*, **9**(12), 3889–3892.

Heinrich, R., Neel, B. G., and Rapoport, T. A. (2002). Mathematical Models of Protein Kinase Signal Transduction. *Molecular Cell*, **9**, 957–970.

Hetherington, A. M. (2001). Guard Cell Signaling. *Cell*, **107**(6), 711 – 714.

Hetherington, A. M. and Woodward, F. I. (2003). The role of stomata in sensing and driving environmental change. *Nature*, **424**(6951), 901–908.

Higham, D. J. (2008). Modeling and Simulating Chemical Reactions. *SIAM Review*, **50**(2), 347–368.

Höfer, T. (1999). Model of Intercellular Calcium Oscillations in Hepatocytes: Synchronization of Heterogeneous Cells. *Biophysical Journal*, **77**(3), 1244–1256.

Höfer, T., Nathansen, H., Löhning, M., Radbruch, A., and Heinrich, R. (2002). GATA-3 transcriptional imprinting in Th2 lymphocytes: a mathematical model. *Proc Natl Acad Sci USA*, **99**(14), 9364–9368.

Hooshangi, S., Thiberge, S., and Weiss, R. (2005). Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proc Natl Acad Sci USA*, **102**(10), 3581–3586.

Howison, S. (2003). *Practical Applied Mathematics: Modelling, Analysis, Approximation*. Cambridge Series in Applied Mathematics. Cambridge University Press.

Huang, C. Y. and Ferrell, J. E. (1996). Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proc Natl Acad Sci USA*, **93**(19), 10078–10083.

Hubbard, K. E., Nishimura, N., Hitomi, K., Getzoff, E. D., and Schroeder, J. I. (2010). Early abscisic acid signal transduction mechanisms: newly discovered components and newly emerging questions. *Genes Dev*, **24**(16), 1695–1708.

Huberman, B. A., Romero, D. M., and Wu, F. (2008). Social networks that matter: Twitter under the microscope. *arXiv:0812.1045*.

Iliopoulos, D., Hintze, A., and Adami, C. (2010). Critical Dynamics in the Evolution of Stochastic Strategies for the Iterated Prisoner's Dilemma. *PLoS Comput Biol*, **6**(10), e1000948.

Ingram, P., Stumpf, M., and Stark, J. (2006). Network motifs: structure does not determine function. *BMC Genomics*, **7**(1), 108.

Islam, M. M., Hossain, M. A., Jannat, R., Munemasa, S., Nakamura, Y., Mori, I. C., and Murata, Y. (2010). Cytosolic Alkalization and Cytosolic Calcium Oscillation in Arabidopsis Guard Cells Response to ABA and MeJA. *Plant and Cell Physiology*, **51**(10), 1721–1730.

Jahan, M. S., Ogawa, K., Nakamura, Y., Shimoishi, Y., Mori, I. C., and Murata, Y. (2008). Deficient Glutathione in Guard Cells Facilitates Abscisic Acid-Induced Stomatal Closure but Does Not Affect Light-Induced Stomatal Opening. *Bioscience, Biotechnology, and Biochemistry*, **72**(10), 2795–2798.

Jammes, F., Song, C., Shin, D., Munemasa, S., Takeda, K., Gu, D., Cho, D., Lee, S., Giordo, R., Sritubtim, S., Leonhardt, N., Ellis, B. E., Murata, Y., and Kwak, J. M. (2009). MAP kinases MPK9 and MPK12 are preferentially expressed in guard cells and positively regulate ROS-mediated ABA signaling. *Proc Natl Acad Sci USA*, **106**(48), 20520–20525.

Jones, R. J. and Mansfield, T. A. (1970). Suppression of Stomatal Opening in Leaves Treated with Abscisic Acid. *Journal of Experimental Botany*, **21**(3), 714–719.

Kang, J., Hwang, J.-U., Lee, M., Kim, Y.-Y., Assmann, S. M., Martinoia, E., and Lee, Y. (2010). PDR-type ABC transporter mediates cellular uptake of the phytohormone abscisic acid. *Proc Natl Acad Sci USA*, **107**(5), 2355–2360.

Kauffman, S. (1969). Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of Theoretical Biology*, **22**(3), 437 – 467.

Ke, R. (2010). *Mathematical Modelling of Ion Regulation in Fungi*. Ph.D. thesis, Imperial College London.

Keener, J. and Sneyd, J. (1998). *Mathematical Physiology*. Springer.

Kholodenko, B. N. (2000). Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades. *Eur J Biochem*, **267**(6), 1583–1588.

Kim, J.-R., Kim, J., Kwon, Y.-K., Lee, H.-Y., Heslop-Harrison, P., and Cho, K.-H. (2011). Reduction of Complex Signaling Networks to a Representative Kernel. *Sci. Signal.*, **4**(175), ra35.

Kim, T.-H., Bhmer, M., Hu, H., Nishimura, N., and Schroeder, J. I. (2010). Guard cell signal transduction network: advances in understanding abscisic acid, CO2, and Ca2+ signaling. *Annual Review of Plant Biology*, **61**, 561–591.

Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, **220**(4598), 671–680.

Klessig, D. F., Durner, J., Noad, R., Navarre, D. A., Wendehenne, D., Kumar, D., Zhou, J. M., Shah, J., Zhang, S., Kachroo, P., Trifa, Y., Pontier, D., Lam, E., and Silva, H. (2000). Nitric oxide and salicylic acid signaling in plant defense. *Proc Natl Acad Sci USA*, **97**(16), 8849–8855.

Klotz, L.-O., Schroeder, P., and Sies, H. (2002). Peroxynitrite signaling: receptor tyrosine kinases and activation of stress-responsive pathways. *Free Radic Biol Med*, **33**(6), 737–743.

Köhler, B., Hills, A., and Blatt, M. R. (2003). Control of Guard Cell Ion Channels by Hydrogen Peroxide and Abscisic Acid Indicates Their Action through Alternate Signaling Pathways. *Plant Physiology*, **131**(2), 385–388.

Koiwai, H., Nakaminami, K., Seo, M., Mitsuhashi, W., Toyomasu, T., and Koshiba, T. (2004). Tissue-Specific Localization of an Abscisic Acid Biosynthetic Enzyme, AAO3, in Arabidopsis. *Plant Physiology*, **134**(4), 1697–1707.

Kreyszig, E. (2006). *Advanced Engineering Mathematics*. John Wiley & Sons, 9 edition.

Kuznetsov, Y. A. (1998). *Elements of applied bifurcation theory*. Applied mathematical sciences. Springer.

Kwak, J. M., C.Mori, I., Pei, Z.-M., Leonhardt, N., Torres, M. A., L.Dangl, J., E.Bloom, R., Bodde, S., D.G.Jones, J., and I.Schroeder, J. (2003). NADPH oxidase AtrbohD and AtrbohF genes function in ROS-dependent ABA signaling in Arabidopsis. *The EMBO Journal*, **22**, 2623–2633.

Kwak, J. M., Nguyen, V., and Schroeder, J. I. (2006). The role of reactive oxygen species in hormonal responses. *Plant Physiology*, **141**(2), 323–329.

Kwak, J. M., Mäser, P., and Schroeder, J. (2008). The Clickable Guard Cell, Version II: Interactive Model of Guard Cell Signal Transduction Mechanisms and Pathways. In *The Arabidopsis Book*, pages 1–17. The American Society of Plant Biologists.

Lamotte, O., Courtois, C., Barnavon, L., Pugin, A., and Wendehenne, D. (2005). Nitric oxide in plants: the biosynthesis and cell signalling properties of a fascinating molecule. *Planta*, **221**(1), 1–4.

Lawson, C. and Hanson, R. (1995). *Solving least squares problems*. Classics in applied mathematics. SIAM.

Lawson, T. (2009). Guard Cell Photosynthesis and Stomatal Function. *New Phytologist*, **181**(1), 13–34.

Levchenko, V., Konrad, K. R., Dietrich, P., Roelfsema, M. R. G., and Hedrich, R. (2005). Cytosolic abscisic acid activates guard cell anion channels without preceding Ca2+ signals. *Proc Natl Acad Sci USA*, **102**(11), 4203–4208.

Levitt, L. K., Stein, D. B., and Rubinstein, B. (1987). Promotion of Stomatal Opening by Indoleacetic Acid and Ethrel in Epidermal Strips of Vicia faba L. *Plant Physiology*, **85**(2), 318–321.

Li, J., Wang, X.-Q., Watson, M. B., and Assmann, S. M. (2000). Regulation of Abscisic Acid-Induced Stomatal Closure and Anion Channels by Guard Cell AAPK Kinase. *Science*, **287**(5451), 300–303.

Li, S., Assmann, S. M., and Albert, R. (2006). Predicting essential components of signal transduction networks: A dynamic model of guard cell abscisic acid signaling. *PLoS Biol*, **4**(10), e312.

Libourel, I. G. L., Bethke, P. C., Michele, R. D., and Jones, R. L. (2006). Nitric oxide gas stimulates germination of dormant Arabidopsis seeds: use of a flow-through apparatus for delivery of nitric oxide. *Planta*, **223**(4), 813–820.

Liu, J., Liu, G., Hou, L., *et al.* (2010a). Ethylene-induced nitric oxide production and stomatal closure in Arabidopsis thaliana depending on changes in cytosolic pH. *Chinese Sci Bull*, **55**, 2403–2409.

Liu, J., Mehdi, S., Topping, J., Tarkowski, P., and Lindsey, K. (2010b). Modelling and experimental analysis of hormonal crosstalk in Arabidopsis. *Molecular Systems Biology*, **6**, 373.

Locke, J., Millar, A., and Turner, M. (2005). Modelling genetic networks with noisy and varied experimental data: the circadian clock in Arabidopsis thaliana. *Journal of Theoretical Biology*, **234**(3), 383 – 393.

Lucius, A. L., Maluf, N. K., Fischer, C. J., and Lohman, T. M. (2003). General Methods for Analysis of Sequential n-step Kinetic Mechanisms: Application to Single Turnover Kinetics of Helicase-Catalyzed DNA Unwinding. *Biophysical Journal*, **85**(4), 2224–2239.

Ma, Y., Szostkiewicz, I., Korte, A., Moes, D., Yang, Y., Christmann, A., and Grill, E. (2009). Regulators of PP2C Phosphatase Activity Function as Abscisic Acid Sensors. *Science*, **324**(5930), 1064–1068.

MacRobbie, E. A. (1998). Signal transduction and ion channels in guard cells. *Philos Trans R Soc Lond B Biol Sci*, **353**(1374), 1475–1488.

Maini, P. K., Burke, M. A., and Murray, J. D. (1991). On the Quasi-Steady-State Assumption Applied to Michaelis-Menten and Suicide Substrate Reactions with Diffusion. *Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences*, **337**(1646), 299–306.

Marchi, M., D'Antoni, A., Formentini, I., Parra, R., Brambilla, R., Ratto, G. M., and Costa, M. (2008). The n-terminal domain of erk1 accounts for the functional differences with erk2. *PLoS One*, **3**(12), e3873.

Marks, F., Klingmüller, U., and Müller-Decker, K. (2009). *Cellular signal processing: an introduction to the molecular mechanisms of signal transduction*. Garland Science.

Matilla, A. J. (2000). Ethylene in seed formation and germination. *Seed Science Research*, **10**(02), 111–126.

Mayo, A. E., Setty, Y., Shavit, S., Zaslaver, A., and Alon, U. (2006). Plasticity of the cis-regulatory input function of a gene. *PLoS Biol*, **4**(4), e45.

Mazzucchelli, C., Vantaggiato, C., Ciamei, A., Fasano, S., Pakhotin, P., Krezel, W., Welzl, H., Wolfer, D. P., Pagès, G., Valverde, O., Marowsky, A., Porrazzo, A., Orban, P. C., Maldonado, R., Ehrengruber, M. U., Cestari, V., Lipp, H.-P., Chapman, P. F., Pouysségur, J., and Brambilla, R. (2002). Knockout of erk1 map kinase enhances synaptic plasticity in the striatum and facilitates striatal-mediated learning and memory. *Neuron*, **34**(5), 807–820.

McAinsh, Brownlee, H. M. R. M. (1990). Abscisic acid-induced elevation of guard cell cytosolic Ca2+ precedes stomatal closure. *Nature*, **343**(6254), 186–188.

McAinsh, M. R., Clayton, H., Mansfield, T. A., and Hetherington, A. M. (1996). Changes in Stomatal Behavior and Guard Cell Cytosolic Free Calcium in Response to Oxidative Stress. *Plant Physiology*, **111**(4), 1031–1042.

McCourt, P. and Creelman, R. (2008). The ABA receptors - we report you decide. *Current Opinion in Plant Biology*, **11**(5), 474 – 478. Cell Signalling and Gene Regulation - Edited by Jason Reed and Bonnie Bartel.

McCue, S., Bartsch, T., Dyson, R., Beguerisse-Díaz, M., and Jensen, O. (2009). Modelling cell separation during plant organ abscission. Technical report, Mathematics in the Plant Sciences Study Group II, University of Nottingham.

McInnis, S. M., Emery, D. C., Porter, R., Desikan, R., Hancock, J. T., and Hiscock, S. J. (2006). The role of stigma peroxidases in flowering plants: insights from further characterization of a stigma-specific peroxidase (ssp) from senecio squalidus (asteraceae). *J Exp Bot*, **57**(8), 1835–1846.

Meinhard, M. and Grill, E. (2001). Hydrogen peroxide is a regulator of ABI1, a protein phosphatase 2C from Arabidopsis. *FEBS Letters*, **508**(3), 443 – 446.

Melcher, K., Ng, L.-M., Zhou, X. E., Soon, F.-F., Xu, Y., Suino-Powell, K. M., Park, S.-Y., Weiner, J. J., Fujii, H., Chinnusamy, V., Kovach, A., Li, J., Wang, Y., Li, J., Peterson, F. C., Jensen, D. R., Yong, E.-L., Volkman, B. F., Cutler, S. R., Zhu, J.-K., and Xu, H. E. (2009). A gate-latch-lock mechanism for hormone signalling by abscisic acid receptors. *Nature*, **462**(7273), 602–608.

Melcher, K., Zhou, X. E., and Xu, H. E. (2010). Thirsty plants and beyond: structural mechanisms of abscisic acid perception and signaling. *Curr Opin Struct Biol*, **20**(6), 722–729.

Melotto, M., Underwood, W., Koczan, J., Nomura, K., and He, S. Y. (2006). Plant stomata function in innate immunity against bacterial invasion. *Cell*, **126**(5), 969–980.

Merlot, S., Gosti, F., Guerrier, D., Vavasseur, A., and Giraudat, J. (2001). The ABI1 and ABI2 protein phosphatases 2C act in a negative feedback regulatory loop of the abscisic acid signalling pathway. *The Plant Journal*, **25**(3), 295–303.

Meyers, L. A., Newman, M. E. J., and Pourbohloul, B. (2006). Predicting epidemics on directed contact networks. *Journal of Theoretical Biology*, **240**(3), 400–418.

Miao, Y., Lv, D., Wang, P., Wang, X.-C., Chen, J., Miao, C., and Song, C.-P. (2006). An Arabidopsis Glutathione Peroxidase Functions as Both a Redox Transducer and a Scavenger in Abscisic Acid and Drought Stress Responses. *Plant Cell*, **18**(10), 2749–2766.

Michaelis, L. and Menten, M. L. (1913). Die kinetik der Invertinwirkung. *Biochem. Z*, **49**(333-369), 352.

Miedema, H. and Assmann, S. M. (1996). A membrane-delimited effect of internal pH on the K+ outward rectifier of Vicia faba guard cells. *Journal of Membrane Biology*, **154**(3), 227–237.

Milner, R., Parrow, J., and Walker, D. (1992). A calculus of mobile processes, I. *Information and Computation*, **100**(1), 1 – 40.

Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill, New York.

Miyazono, K.-i., Miyakawa, T., Sawano, Y., Kubota, K., Kang, H.-J., Asano, A., Miyauchi, Y., Takahashi, M., Zhi, Y., Fujita, Y., Yoshida, T., Kodaira, K.-S., Yamaguchi-Shinozaki, K., and Tanokura, M. (2009). Structural basis of abscisic acid signalling. *Nature*, **462**(7273), 609–614.

Mohr, H. and Schopfer, P. (1995). *Plant physiology*. Springer.

Moles, C. G., Mendes, P., and Banga, J. R. (2003). Parameter Estimation in Biochemical Pathways: A Comparison of Global Optimization Methods. *Genome Research*, **13**(11), 2467–2474.

Monk, N. A. M. (2003). Oscillatory expression of Hes1, p53, and NF-kappaB driven by transcriptional time delays. *Current Biology*, **13**(16), 1409–1413.

Mood, A. M., Graybill, F. A., and Boes, D. C. (1974). *Introduction to the theory of statistics*. McGraw-Hill.

Mucha, P. J., Richardson, T., Macon, K., Porter, M. A., and Onnela, J.-P. (2010). Community Structure in Time-Dependent, Multiscale, and Multiplex Networks. *Science*, **328**(5980), 876–878.

Müller, A. H. and Hansson, M. (2009). The Barley Magnesium Chelatase 150-kD Subunit Is Not an Abscisic Acid Receptor. *Plant Physiology*, **150**(1), 157–166.

Muller, M. and Ram, P. T. (2010). Systems biology of the mapk1,2 network. In S. Choi, editor, *Systems Biology for Signaling Networks*, volume 1 of *Systems Biology*, pages 455–489. Springer New York.

Murata, Y., Pei, Z.-M., Mori, I. C., and Schroeder, J. (2001). Abscisic Acid Activation of Plasma Membrane Ca2+ Channels in Guard Cells Requires Cytosolic NAD(P)H and Is Differentially Disrupted Upstream and Downstream of Reactive Oxygen Species Production in abi1-1 and abi2-1 Protein Phosphatase 2C Mutants. *Plant Cell*, **13**(11), 2513–2523.

Murray, J. (2005). *Mathematical Biology: I. An Introduction*. Springer.

Mustilli, A.-C., Merlot, S., Vavasseur, A., Fenzi, F., and Giraudat, J. (2002). Arabidopsis OST1 Protein Kinase Mediates the Regulation of Stomatal Aperture by Abscisic Acid and Acts Upstream of Reactive Oxygen Species Production. *Plant Cell*, **14**(12), 3089–3099.

Nambara, E. and Marion-Poll, A. (2005). Abscisic Acid Biosynthesis and Catabolism. *Annual Review of Plant Biology*, **56**, 165–185.

Neill, S., Desikan, R., and Hancock, J. (2002). Hydrogen peroxide signalling. *Current Opinion in Plant Biology*, **5**(5), 388 – 395.

Neill, S., Barros, R., Bright, J., Desikan, R., Hancock, J., Harrison, J., Morris, P., Ribeiro, D., and Wilson, I. (2008). Nitric oxide, stomatal closure, and abiotic stress. *Journal of Experimental Botany*, **59**(2), 165–176.

Neill, S. J., Desikan, R., and Hancock, J. T. (2003). Nitric oxide signalling in plants. *New Phytologist*, **159**(1), 11–35.

Nelder, J. A. and Mead, R. (1965). A Simplex Method for Function Minimization. *The Computer Journal*, **7**(4), 308–313.

Newman, M. E. J. (2003). The Structure and Function of Complex Networks. *SIAM Review*, **45**(2), 167–256.

Ng, C. K.-Y., Mcainsh, M. R., Gray, J. E., Hunt, L., Leckie, C. P., Mills, L., and Hetherington, A. M. (2001). Calcium-based signalling systems in guard cells. *New Phytologist*, **151**(1), 109–120.

Nishimura, N., Hitomi, K., Arvai, A. S., Rambo, R. P., Hitomi, C., Cutler, S. R., Schroeder, J. I., and Getzoff, E. D. (2009). Structural mechanism of abscisic acid binding and signaling by dimeric PYR1. *Science*, **326**(5958), 1373–1379.

Nocedal, J. and Wright, S. J. (2006). *Numerical Optimization*. Springer, 2nd edition.

Ogasawara, Y., Kaya, H., Hiraoka, G., Yumoto, F., Kimura, S., Kadota, Y., Hishinuma, H., Senzaki, E., Yamagoe, S., Nagata, K., Nara, M., Suzuki, K., Tanokura, M., and Kuchitsu, K. (2008). Synergistic Activation of the Arabidopsis NADPH Oxidase AtrbohD by Ca2+ and Phosphorylation. *Journal of Biological Chemistry*, **283**(14), 8885–8892.

Onnela, J.-P., Chakraborti, A., Kaski, K., Kertész, J., and Kanto, A. (2003). Dynamics of market correlations: Taxonomy and portfolio analysis. *Physical Review E*, **68**, 056110.

Pagès, G., Gurin, S., Grall, D., Bonino, F., Smith, A., Anjuere, F., Auberger, P., and Pouysségur, J. (1999). Defective thymocyte maturation in p44 map kinase (erk 1) knockout mice. *Science*, **286**(5443), 1374–1377.

Palla, G., Barabási, A.-L., and Vicsek, T. (2007). Quantifying social group evolution. *Nature*, **446**(7136), 664–667.

Pallas, James E., J. and Kays, S. J. (1982). Inhibition of Photosynthesis by Ethylene – A Stomatal Effect . *Plant Physiology*, **70**(2), 598–601.

Pandey, S., Zhang, W., and Assmann, S. M. (2007). Roles of ion channels and transporters in guard cell signal transduction. *FEBS Lett*, **581**(12), 2325–2336.

Pandey, S., Nelson, D. C., and Assmann, S. M. (2009). Two Novel GPCR-Type G Proteins are Abscisic Acid Receptors in Arabidopsis. *Cell*, **136**(1), 136 – 148.

Papachristodoulou, A. and Recht, B. (2007). Determining Interconnections in Chemical Reaction Networks. In *American Control Conference, 2007.*, pages 4872–4877.

Paris, R. B. (2010). Incomplete Gamma and Related Functions. In *NIST Digital Library of Mathematical Functions*.

Park, S.-Y., Fung, P., Nishimura, N., Jensen, D. R., Fujii, H., Zhao, Y., Lumba, S., Santiago, J., Rodrigues, A., Chow, T.-f. F., Alfred, S. E., Bonetta, D., Finkelstein, R., Provart, N. J., Desveaux, D., Rodriguez, P. L., McCourt, P., Zhu, J.-K., Schroeder, J. I., Volkman, B. F., and Cutler, S. R. (2009). Abscisic Acid Inhibits Type 2C Protein Phosphatases via the PYR/PYL Family of START Proteins. *Science*, **324**(5930), 1068–1071.

Pei, Z.-M. and Kuchitsu, K. (2005). Early ABA Signaling Events in Guard Cells. *Journal of Plant Growth Regulation*, **24**(4), 296–307.

Pei, Z. M., Ward, J. M., Harper, J. F., and Schroeder, J. I. (1996). A novel chloride channel in vicia faba guard cell vacuoles activated by the serine/threonine kinase, cdpk. *EMBO J*, **15**(23), 6564–6574.

Pei, Z. M., Kuchitsu, K., Ward, J. M., Schwarz, M., and Schroeder, J. I. (1997). Differential Abscisic Acid Regulation of Guard Cell Slow Anion Channels in Arabidopsis Wild-Type and abi1 and abi2 Mutants. *Plant Cell*, **9**(3), 409–423.

Pei, Z.-M., Baizabal-Aguirre, V. M., Allen, G. J., and Schroeder, J. I. (1998). A transient outward-rectifying K+ channel current down-regulated by cytosolic Ca2+ in Arabidopsis thaliana guard cells. *Proc Natl Acad Sci USA*, **95**(11), 6548–6553.

Pei, Z.-M., Murata, Y., Benning, G., Thomine, S., Klusener, B., Allen, G. J., Grill, E., and Schroeder, J. I. (2000). Calcium channels activated by hydrogen peroxide mediate abscisic acid signalling in guard cells. *Nature*, **406**, 731–734.

Perazzolli, M., Romero-Puertas, M. C., and Delledonne, M. (2006). Modulation of nitric oxide bioactivity by plant haemoglobins. *J Exp Bot*, **57**(3), 479–488.

Pham, J. and Desikan, R. (2009). ROS Signalling in Stomata. In L. A. Rio and A. Puppo, editors, *Reactive Oxygen Species in Plant Signaling*, Signaling and Communication in Plants, pages 55–71. Springer Berlin Heidelberg.

Po, H. N. and Senozan, N. M. (2001). The Henderson-Hasselbalch Equation: Its History and Limitations. *Journal of Chemical Education*, **78**(11), 1499–.

Powell, M. J. D. (1998). Direct Search Algorithms for Optimization Calculations. *Acta Numerica*, **7**(-1), 287–336.

Prajna, S. and Sandberg, H. (2005). On Model Reduction of Polynomial Dynamical Systems. pages 1666–1671.

Rayner, S. A., King, W. J., Comer, R. M., Isaacs, J. D., Hale, G., George, A. J., and Larkin, D. F. (2000). Local bioactive tumour necrosis factor (TNF) in corneal allotransplantation. *Clin Exp Immunol*, **122**(1), 109–116.

Raz, V., Bergervoet, J. H., and Koornneef, M. (2001). Sequential steps for developmental arrest in Arabidopsis seeds. *Development*, **128**(2), 243–252.

Razem, F. A., El-Kereamy, A., Abrams, S. R., and Hill, R. D. (2006). The RNA-binding protein FCA is an abscisic acid receptor. *Nature*, **439**(7074), 290–294.

Razem, F. A., El-Kereamy, A., Abrams, S. R., and Hill, R. D. (2008). Retraction. The RNA-binding protein FCA is an abscisic acid receptor. *Nature*, **456**(7223), 824.

Risk, J. M., Macknight, R. C., and Day, C. L. (2008). FCA does not Bind Abscisic Acid. *Nature*, **456**(7223), E5–E6.

Rockel, P., Strube, F., Rockel, A., Wildt, J., and Kaiser, W. M. (2002). Regulation of nitric oxide (NO) production by plant nitrate reductase in vivo and in vitro. *J Exp Bot*, **53**(366), 103–110.

Runarsson, T. and Yao, X. (2000). Stochastic ranking for constrained evolutionary optimization. *IEEE Transactions on Evolutionary Computation*, **4**(3), 284 –294.

Runarsson, T. and Yao, X. (2005). Search biases in constrained evolutionary optimization. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, **35**(2), 233 –243.

Saadatpour, A., Albert, I., and Albert, R. (2010). Attractor analysis of asynchronous Boolean models of signal transduction networks. *Journal of Theoretical Biology*, **266**, 641–656.

Santiago, J., Dupeux, F., Round, A., Antoni, R., Park, S.-Y., Jamin, M., Cutler, S. R., Rodriguez, P. L., and Mrquez, J. A. (2009). The abscisic acid receptor PYR1 in complex with abscisic acid. *Nature*, **462**(7273), 665–668.

Sauro, H. M. (2009). Network Dynamics. In R. Ireton, K. Montgomery, R. Bumgarner, R. Samudrala, and J. McDermott, editors, *Computational Systems Biology*, volume 541 of *Methods in Molecular Biology*, pages 269–309. Humana Press.

Sauter, A., Davies, W. J., and Hartung, W. (2001). The long-distance abscisic acid signal in the droughted plant: the fate of the hormone on its way from root to shoot. *Journal of Experimental Botany*, **52**(363), 1991–1997.

Schroeder, J. I. and Fang, H. H. (1991). Inward-rectifying k+ channels in guard cells provide a mechanism for low-affinity k+ uptake. *Proc Natl Acad Sci U S A*, **88**(24), 11583–11587.

Schroeder, J. I. and Hagiwara, S. (1989). Cytosolic calcium regulates ion channels in the plasma membrane of vicia faba guard cells. *Nature*, **338**(6214), 427–430.

Schroeder, J. I., Allen, G. J., Hugouvieux, V., Kwak, J. M., and Waner, D. (2001). Guard cell signal transduction. *Annual Review of Plant Physiology and Plant Molecular Biology*, **52**(1), 627–658. PMID: 11337411.

Schwefel, H. (1995). *Evolution and optimum seeking*. Sixth-generation computer technology series. Wiley.

Schwikowski, B., Uetz, P., and Fields, S. (2000). A network of protein-protein interactions in yeast. *Nat Biotechnol*, **18**(12), 1257–1261.

Segel, L. A. and Slemrod, M. (1989). The Quasi-Steady-State Assumption: A Case Study in Perturbation. *SIAM Review*, **31**(3), 446–477.

Seydel, R. (2010). *Practical Bifurcation and Stability Analysis*. Interdisciplinary Applied Mathematics. Springer Verlag.

Shen, Y.-Y., Wang, X.-F., Wu, F.-Q., Du, S.-Y., Cao, Z., Shang, Y., Wang, X.-L., Peng, C.-C., Yu, X.-C., Zhu, S.-Y., Fan, R.-C., Xu, Y.-H., and Zhang, D.-P. (2006). The Mg-Chelatase H Subunit is an Abscisic Acid Receptor. *Nature*, **443**(7113), 823–826.

Siahaan, H. B. (2008). A Balancing Approach to Model Reduction of Polynomial Nonlinear Systems. In *World Congress*, volume 17.

Siegel, R. S., Xue, S., Murata, Y., Yang, Y., Nishimura, N., Wang, A., and Schroeder, J. I. (2009). Calcium elevation-dependent and attenuated resting calcium-dependent abscisic acid induction of stomatal closure and abscisic acid-induced enhancement of calcium sensitivities of s-type anion and inward-rectifying k channels in arabidopsis guard cells. *Plant J*, **59**(2), 207–220.

Sirichandra, C., Wasilewska, A., Vlad, F., Valon, C., and Leung, J. (2009). The guard cell as a single-cell model towards understanding drought tolerance and abscisic acid action. *Journal of Experimental Botany*, **60**(5), 1439–1463.

Sisson, S. A., Fan, Y., and Tanaka, M. M. (2007). Sequential Monte Carlo without likelihoods. *Proc Natl Acad Sci USA*, **104**(6), 1760–1765.

Sokolovski, S. and Blatt, M. R. (2004). Nitric Oxide Block of Outward-Rectifying K+ Channels Indicates Direct Control by Protein Nitrosylation in Guard Cells. *Plant Physiology*, **136**(4), 4275–4284.

Sokolovski, S., Hills, A., Gay, R., Garcia-Mata, C., Lamattina, L., and Blatt, M. R. (2005). Protein phosphorylation is a prerequisite for intracellular Ca2+ release and ion channel control by nitric oxide and abscisic acid in guard cells. *The Plant Journal*, **43**(4), 520–529.

Soni, A., Jenkins, J., and Sundaram, S. (2008). Determination of critical network interactions: an augmented Boolean pseudo-dynamics approach. *IET Systems Biology*, **2**(2), 55–63.

Sontag, E. (2002). For Differential Equations with r Parameters, 2r+1 Experiments Are Enough for Identification. *Journal of Nonlinear Science*, **12**, 553–583.

Stark, J. and Hardy, K. (2003). Chaos: Useful at Last? *Science*, **301**(5637), 1192–1193.

Stark, J., Chan, C., and George, A. J. T. (2007). Oscillations in the immune system. *Immunological Reviews*, **216**(1), 213–231.

Stepanova, A. N. and Alonso, J. M. (2005). Arabidopsis Ethylene Signaling Pathway. *Sci. STKE*, **2005**(276), cm4.

Strelkowa, N. and Barahona, M. (2011). Transient dynamics around unstable periodic orbits in the generalized repressilator model. *Chaos*, **21**(2), 023104.

Stricker, J., Cookson, S., Bennett, M. R., Mather, W. H., Tsimring, L. S., and Hasty, J. (2008). A fast, robust and tunable synthetic gene oscillator. *Nature*, **456**(7221), 516–519.

Strogatz, S. H. (1994). *Nonlinear Dynamics And Chaos. With Applications to Physics, Biology, Chemistry, and Engineering*. Studies in nonlinearity. Perseus Books Group.

Stumpf, M. and Wiuf, C. (2010). A Network Analysis Primer. In M. Stumpf and C. Wiuf, editors, *Statistical and evolutionary analysis of biological networks*, pages 1–15. Imperial College Press.

Suhita, D., Raghavendra, A. S., Kwak, J. M., and Vavasseur, A. (2004). Cytoplasmic Alkalization Precedes Reactive Oxygen Species Production during Methyl Jasmonate- and Abscisic Acid-Induced Stomatal Closure. *Plant Physiology*, **134**(4), 1536–1545.

Süli, E. and Mayers, D. (2003). *Introduction to Numerical Analysis*. Cambridge University Press.

Szallasi, Z., Stelling, J., and Periwal, V. (2006). *System modeling in cell biology: from concepts to nuts and bolts*. Bradford Book. MIT Press.

Taiz, L. and Zeiger, E. (2002). *Plant Physiology*. Sinauer Associates, Inc., 3 edition.

Tan, B.-C., Joseph, L. M., Deng, W.-T., Liu, L., Li, Q.-B., Cline, K., and McCarty, D. R. (2003). Molecular characterization of the Arabidopsis 9-cis epoxycarotenoid dioxygenase gene family. *The Plant Journal*, **35**(1), 44–56.

Tanaka, Y., Sano, T., Tamaoki, M., Nakajima, N., Kondo, N., and Hasezawa, S. (2005). Ethylene inhibits abscisic acid-induced stomatal closure in Arabidopsis. *Plant Physiology*, **138**(4), 2337–2343.

Tanaka, Y., Sano, T., Tamaoki, M., Nakajima, N., Kondo, N., and Hasezawa, S. (2006). Cytokinin and auxin inhibit abscisic acid-induced stomatal closure by enhancing ethylene production in Arabidopsis. *Journal of Experimental Botany*, **57**(10), 2259–2266.

Taylor, H. and Karlin, S. (1998). *An introduction to stochastic modeling*. Academic Press.

Thoday, D. (1952). Turgor Pressure and Wall Pressure. *Ann Bot*, **16**(2), 129–131.

Toni, T. and Stumpf, M. P. H. (2010). Simulation-based model selection for dynamical systems in systems and population biology. *Bioinformatics*, **26**(1), 104–110.

Toni, T., Welch, D., Strelkowa, N., Ipsen, A., and Stumpf, M. P. (2009). Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of The Royal Society Interface*, **6**(31), 187–202.

Traud, A. L., Kelsic, E. D., Mucha, P. J., and Porter, M. A. (2011). Comparing Community Structure to Characteristics in Online Collegiate Social Networks. *SIAM Review*, **53**(3), 526–543.

Tyson, J. J., Chen, K. C., and Novak, B. (2003). Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Current Opinion in Cell Biology*, **15**(2), 221 – 231.

Van Loan, C. F. (1976). Computing Integrals Involving the Matrix Exponential. Technical report, Cornell University, Ithaca, NY, USA.

Vlad, F., Rubio, S., Rodrigues, A., Sirichandra, C., Belin, C., Robert, N., Leung, J., Rodriguez, P. L., Laurire, C., and Merlot, S. (2009). Protein phosphatases 2C regulate the activation of the Snf1-related kinase OST1 by abscisic acid in Arabidopsis. *Plant Cell*, **21**(10), 3170–3184.

Waage, P. and Gulberg, C. M. (1986). Studies concerning affinity. *Journal of Chemical Education*, **63**(12), 1044. H.I. Abrash, translator.

Wang, B. (2010). *Design and Functional Assembly of Synthetic Biological Parts and Devices*. Ph.D. thesis, Imperial College London.

Wang, B., Kitney, R. I., Joly, N., and Buck, M. (2011). Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology. *Nat Commun*, **2**, 508.

Wang, P. and Song, C.-P. (2008). Guard-cell signalling for hydrogen peroxide and abscisic acid. *New Phytologist*, **178**(4), 703–718.

Wang, X.-Q., Ullah, H., Jones, A. M., and Assmann, S. M. (2001). G Protein Regulation of Ion Channels and Abscisic Acid Signaling in Arabidopsis Guard Cells. *Science*, **292**(5524), 2070–2072.

Wang, Y., Liu, C., Li, K., Sun, F., Hu, H., Li, X., Zhao, Y., Han, C., Zhang, W., Duan, Y., Liu, M., and Li, X. (2007). Arabidopsis EIN2 modulates stress response through abscisic acid response pathway. *Plant Molecular Biology*, **64**(6), 633–644.

Ward, J. M. and Schroeder, J. I. (1994). Calcium-activated k+ channels and calcium-induced calcium release by slow vacuolar ion channels in guard cell vacuoles implicated in the control of stomatal closure. *Plant Cell*, **6**(5), 669–683.

Wasilewska, A., Vlad, F., Sirichandra, C., Redko, Y., Jammes, F., Valon, C., Frey, N. F. d., and Leung, J. (2008). An Update on Abscisic Acid Signaling in Plants and More ... *Molecular Plant*, **1**(2), 198–217.

Wasserman, S. and Faust, K. (1994). *Social network analysis: methods and applications*. Structural analysis in the social sciences. Cambridge University Press.

Weinberg, R. (2007). *The biology of cancer*. Garland Science.

Weiner, J. J., Peterson, F. C., Volkman, B. F., and Cutler, S. R. (2010). Structural and functional insights into core ABA signaling. *Current Opinion in Plant Biology*, **13**(5), 495 – 502.

Wilkinson, D. J. (2009). Stochastic modelling for quantitative description of heterogeneous biological systems. *Nat Rev Genet*, **10**(2), 122–133.

Wilkinson, S. and Davies, W. J. (2009a). Drought, ozone, ABA and ethylene: new insights from cell to plant to community. *Plant, Cell & Environment*, **33**, 510–525.

Wilkinson, S. and Davies, W. J. (2009b). Ozone suppresses soil drying- and abscisic acid (ABA)-induced stomatal closure via an ethylene-dependent mechanism. *Plant Cell Environ*, **32**(8), 949–959.

Wilson, I. D., Neill, S. J., and Hancock, J. T. (2008). Nitric oxide synthesis and signalling in plants. *Plant, Cell & Environment*, **31**(5), 622–631.

Yang, Y., Costa, A., Leonhardt, N., Siegel, R. S., and Schroeder, J. I. (2008). Isolation of a strong arabidopsis guard cell promoter and its potential as a research tool. *Plant Methods*, **4**, 6.

Yates, A., Chan, C. C. W., Callard, R. E., George, A. J. T., and Stark, J. (2001). An approach to modelling in immunology. *Brief Bioinform*, **2**(3), 245–257.

Yi, S. and Ulsoy, A. (2006). Solution of a system of linear delay differential equations using the matrix Lambert function. In *American Control Conference, 2006*, page 6 pp.

Yoshida, R., Umezawa, T., Mizoguchi, T., Takahashi, S., Takahashi, F., and Shinozaki, K. (2006). The regulatory domain of SRK2E/OST1/SnRK2.6 interacts with ABI1 and integrates abscisic acid (ABA) and osmotic stress signals controlling stomatal closure in Arabidopsis. *Journal of Biological Chemistry*, **281**(8), 5310–5318.

Yoshii, H. and Imaseki, H. (1981). Biosynthesis of auxin-induced ethylene. effects of indole-3-acetic acid, benzyladenine and abscisic acid on endogenous levels of 1-aminocyclopropane-l-carboxylic acid (acc) and acc synthase. *Plant and Cell Physiology*, **22**(3), 369–379.

Young, J. J., Mehta, S., Israelsson, M., Godoski, J., Grill, E., and Schroeder, J. I. (2006). CO2 signaling in guard cells: Calcium sensitivity response modulation, a Ca2+-independent phase, and CO2 insensitivity of the gca2 mutant. *Proc Natl Acad Sci USA*, **103**(19), 7506–7511.

Yun, B.-W., Feechan, A., Yin, M., Saidi, N. B. B., Bihan, T. L., Yu, M., Moore, J. W., Kang, J.-G., Kwon, E., Spoel, S. H., Pallas, J. A., and Loake, G. J. (2011). S-nitrosylation of NADPH oxidase regulates cell death in plant immunity. *Nature*, **478**(7368), 264–268.

Zafra, A., Rodríguez-García, M. I., and de Dios Alché, J. (2010). Cellular localization of ros and no in olive reproductive tissues during flower development. *BMC Plant Biol*, **10**, 36.

Zeng, W., Melotto, M., and He, S. Y. (2010). Plant stomata: a checkpoint of host immunity and pathogen virulence. *Current Opinion in Biotechnology*, **21**(5), 599–603.

Zhang, S. and Klessig, D. F. (2001). MAPK cascades in plant defense signaling. *Trends Plant Sci*, **6**(11), 520–527.

Zhang, X., Dong, F. C., Gao, J. F., and Song, C. P. (2001a). Hydrogen peroxide-induced changes in intracellular pH of guard cells precede stomatal closure. *Cell Research*, **11**(1), 37–43.

Zhang, X., Zhang, L., Dong, F., Gao, J., Galbraith, D. W., and Song, C.-P. (2001b). Hydrogen Peroxide Is Involved in Abscisic Acid-Induced Stomatal Closure in Vicia faba. *Plant Physiology*, **126**(4), 1438–1448.

Zhang, X., Miao, Y. C., An, G. Y., Zhou, Y., Shangguan, Z. P., Gao, J. F., and Song, C. P. (2001c). K+ channels inhibited by hydrogen peroxide mediate abscisic acid signaling in Vicia guard cells. *Cell Research*, **11**(3), 195–202.

Zi, Z. and Klipp, E. (2006). SBML-PET: a Systems Biology Markup Language-based parameter estimation tool. *Bioinformatics*, **22**(21), 2704–2705.