FIXED NITROGEN DYNAMICS AND HETEROCYST PATTERNING IN FILAMENTOUS HETEROCYSTOUS CYANOBACTERIA

by

Aidan I Brown

Submitted in partial fulfillment of the
requirements for the degree of
Master of Science

at

Dalhousie University
Halifax, Nova Scotia
August 2012

DALHOUSIE UNIVERSITY

DEPARTMENT OF PHYSICS AND ATMOSPHERIC SCIENCE

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled "FIXED NITROGEN DYNAMICS AND HETEROCYST PATTERNING IN FILAMENTOUS HETEROCYSTOUS CYANOBACTERIA" by Aidan I Brown in partial fulfillment of the requirements for the degree of Master of Science.

Dated: August 10, 2012

Supervisor: _____

Readers: _____

_____

DALHOUSIE UNIVERSITY

DATE: August 10, 2012

AUTHOR: Aidan I Brown

TITLE: FIXED NITROGEN DYNAMICS AND HETEROCYST
PATTERNING IN FILAMENTOUS HETEROCYSTOUS
CYANOBACTERIA

DEPARTMENT OR SCHOOL: Department of Physics and Atmospheric Science

DEGREE: M.Sc.   CONVOCATION: October   YEAR: 2012

_____
Signature of Author

Table of Contents

List of Figures

Abstract

Cyanobacteria are prokaryotes that can grow photoautotrophically using oxygenic photosynthesis. Some filamentous cyanobacteria in media with insufficient fixed nitrogen develop a regular pattern of heterocyst cells that fix nitrogen for the remaining vegetative cells. We have built an integrated computational model of fixed nitrogen transport and cell growth for filamentous cyanobacteria. With our model, two qualitatively different experimentally observed nitrogen distributions between a pair of heterocysts are reconciled. By adding dynamic heterocyst placement into our model, we can optimize heterocyst frequency with respect to growth. Further introduction of modest leakage leads to distinct growth rates between different heterocyst placement strategies. A local placement strategy yields maximal growth and steady state heterocyst spacings similar to those observed experimentally. Adding more realistic fixed nitrogen storage based heterocyst commitment together with lateral inhibition to the model allows us to address initial heterocyst commitment and qualitatively reproduces many aspects of heterocyst differentiation. We also investigate patterns of starving cells and correlations of fixed nitrogen in filaments without heterocysts. We find percolation transitions in both spatial one dimensional patterns and space-time two dimensional patterns.

List of Abbreviations and Symbols Used

2-OG: 2-oxoglutarate

efN: External fixed nitrogen

fN: fixed nitrogen

FRAP: Fluorescence recovery after photobleaching

GFP: Green fluorescent protein

Mch: Multiple contiguous heterocysts

PCC: Pasteur Culture Collection

WT: Wild type

Acknowledgements

I would like to thank all the people who have helped and inspired me during my research and work on this thesis. I would also like to thank anyone who correctly reminded me that maybe I should be doing work. I would like to thank airports with free wifi and friends with cellphones that can send email for helping me to remember things I otherwise would have likely forgotten.

I first and foremost want to thank my supervisor, Andrew Rutenberg, for his encouragement, guidance, and support during my research, the results of which reside inside and outside of this thesis. I would also like to thank him for putting up with me when I did not listen to his advice, answered emails late or not at all, for explaining things a third or fourth time when I was lost, and for telling me when it was time to figure things out by myself, among other things. I would really like to thank him for his relentless constructive criticism and for telling me to 'own it' and other ways of saying to always try hard and put my best into everything.

I would like to thank my parents for completely supporting my decision to go move to Halifax and loving and supporting me the entire time. I would like to also thank my parents for politely asking about my work and then politely nodding through what were probably poor explanations.

Chapter 1

Introduction

1.1   Motivation

Since the work of Fogg in the 1940's [1, 2], scientists have been trying to understand heterocysts in filamentous cyanobacteria, how they produce and are affected by fixed nitrogen, and how the filament controls its differentiation pattern. The initiation and maintenance of a pattern of heterocysts within filaments is a simple and compelling example of both cell differentiation and pattern formation.
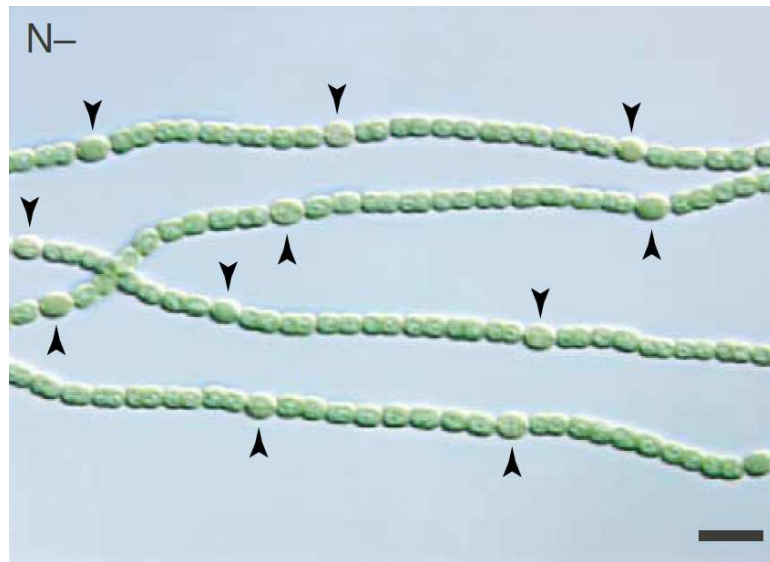
Figure 1.1: Heterocyst pattern in the filamentous cyanobacterium *Anabaena* PCC 7120. Heterocysts are denoted by arrows. N- indicates that the filament was grown in a medium without fixed nitrogen. Figure from Golden and Yoon [3].

Physical pattern formation has long been studied [4], and more recently so has pattern formation in biological systems [5]. Cyanobacterial heterocyst patterns have been studied quantitatively only recently [6], with recent quantitative measurements also involving involving transport [7] and the nitrogen distribution [8], among other measurements.

These quantitative measurements have made the theoretical and computational study of heterocyst differentiation much more appealing because there are now results to compare to and a more quantitatively literate community to speak to.

Of course, in addition to available quantitative measurements, there are other reasons for a physical study of heterocyst differentiation. The primary one for me is the apparent lack of a broken symmetry. In a very wide range of systems, symmetry is broken by an outside influence. In the Ising model of magnetic systems, the symmetry is broken by an external magnetic field [9], and in *Drosophila* embryos, maternal determinants set up the anterior-posterior axis for the process that eventually localizes and orients the mature segments of the fly [10]. In cyanobacteria there is no known outside influence which chooses which cells will differentiate. I want to understand how the filament determines which cells will become heterocysts.

1.2   Hypothesis

We hypothesize that much of the phenomena associated with heterocyst differentiation, including the symmetry breaking behind the formation of the pattern, can be explained with simple, physical rules. We want to discover, understand, and, as much as possible, quantify these rules. Equally important, we want to understand how the collective effect of these rules leads to a pattern of differentiated cells, and how the aspects of the pattern quantitatively depend on the rules. Almost all of the literature on the topic of heterocyst differentiation focuses on the genetic and biochemical details of heterocysts and cyanobacteria, and our understanding will necessarily be in the context of this large body of work.

Our work will focus on the areas that have been quantitatively investigated experimentally. We will investigate how the nitrogen distribution is affected by transport and growth, in the context of relevant studies. We will go beyond the existing investigation of the wild type pattern, and focus on the pattern in growing systems. We will also examine the initiation of the heterocyst pattern and the symmetry breaking that must take place. Finally, we will understand a system of starving cells in the context of traditional physical measurements.

Also, to develop a full model that includes detailed genetic effects (which we have not done), you must first develop a physical model of growth and metabolite exchange.

As such, this thesis forms a strong foundation for further work.

1.3   Outline

The purpose of Chapter 2 is to provide a background to understand the work of Chapters 3-5, and also to provide context for the work. It introduces cyanobacteria, heterocysts, and the heterocyst pattern, and then goes through much of the information in the literature relevant to our investigation of the pattern. Chapter 6 is significantly different than the previous three and the relevant background for that chapter is provided at its beginning.

Chapters 3-6 all use a similar model, but one that evolves between chapters as we test it against experimental data. The chapters investigate different phenomena and the models were chosen to be as simple as possible for those phenomena.

Chapter 3 focuses on the fixed nitrogen distribution in the filament. The disagreement between the experimentally determined fixed nitrogen distributions of Wolk *et al* [11] and Popa *et al* [8] is resolved and cytoplasmic and periplasmic transport are discussed.

Chapter 4 looks at steady state growth of a filament with heterocysts that has ongoing heterocyst placement. Comparisons between filaments with different heterocyst placement strategies and external fixed nitrogen concentrations are made using the filament growth rate and the heterocyst patterns in the filament.

Chapter 5 investigates initial pattern formation in filaments deprived of external fixed nitrogen. Storage of fixed nitrogen and lateral inhibition of differentiation are introduced to the model. Several quantities produced by simulation are compared to experimental results from the literature.

Chapter 6 presents the cluster size distributions of starving cells in filaments without heterocysts. These are compared to percolation in one and two dimensions.

Chapter 2

Background

2.1 Cyanobacteria

Cyanobacteria, often known as blue-green algae, are Gram-negative bacteria ranging in diameter from less than one micrometre to greater than 100 micrometres [12]. Cyanobacteria were present in both unicellular and filamentous forms at least 3.5 billion years ago [13, 14]. Cyanobacteria are significant to the history of the earth because they were the first organisms to evolve oxygenic photosynthesis [15] between 2.78 and 3.7 billion years ago [16]. As early photosynthesizers cyanobacteria oxygenated the atmosphere between 2.5 and 2.2 billion years ago [16]. This oxygenation of the atmosphere is hypothesized to have caused a global glaciation (through reduced methane levels) [16], the retreat of many anaerobic life forms due to their decreased viability in the presence of oxygen, and also encouraged the development of the oxygen-requiring respiratory electron transport chain central to the metabolism of much of eukaryotic life today [17].

Cyanobacteria can be found in both freshwater and marine environments as well as on top of and in the soil where they can increase nitrogen content and are important to agriculture [18]. They form symbiotic relationships with other bacteria as well as fungi, lower and higher plants, and invertebrate animals [18]. Cyanobacteria are also thought to be the ancestors of chloroplasts in plants, which were originally endosymbionts [18].

At the cellular level cyanobacteria exist as individual cells, unbranched filaments, or branched filaments [18]. I will focus on unbranched filamentous cyanobacteria, which have several cell types. The cell types allow the cyanobacterium to withstand unfavourable conditions [19]. The first and most common are vegetative cells whose primary function is to perform photosynthesis to provide fixed carbon to the organism [15]. Akinetes are spores which are larger than vegetative cells and only differentiate under adverse conditions [15, 19]. Hormogonia are short, motile filaments composed of
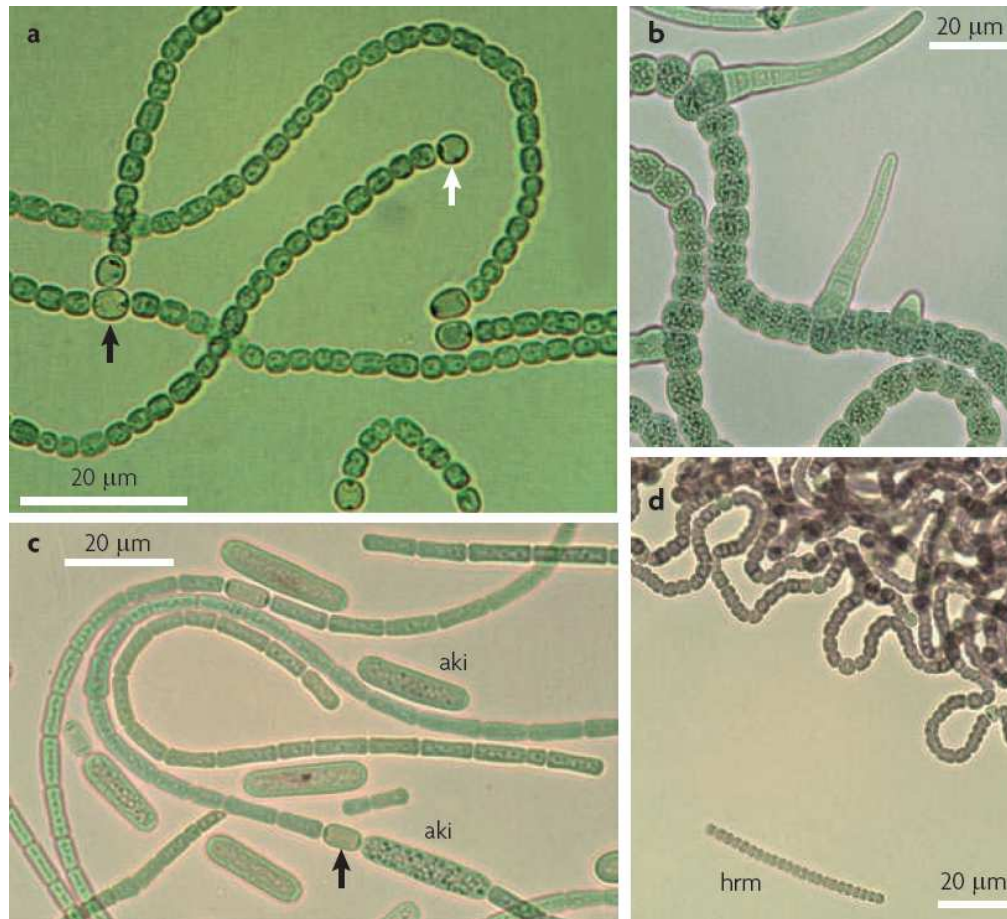
Figure 2.1: (a) Filaments of *Anabaena* PCC 7120 grown on N_2 as the nitrogen source. An intercalary heterocyst is indicated by the black arrow, and a terminal heterocyst is indicated by the white arrow. (b) Branched filaments of *Fischerella muscicola*. (c) N_2 grown filaments of *Anabaena cylindrica*. A heterocyst is indicated by the black arrow and akinetes indicated by 'aki'. (d) N_2 grown *Nostoc* PCC 9203 showing filaments as well as one hormogonium, indicated by 'hrm'. Figure from Flores and Herrero [15].

small cells that are derived from cell division without increase in biomass; hormogonia also only differentiate under adverse conditions [15, 19]. I will focus on filamentous heterocystous cyanobacteria that do not differentiate either akinetes or hormogonia. These filaments consist of vegetative cells and occasional heterocyst cells.

Nitrogen in the atmosphere must be 'fixed' into ammonia and then changed into other forms to be used by life because dinitrogen can otherwise not be metabolized.

The enzyme responsible for nitrogen fixation is nitrogenase, which is poisoned by oxygen [20]. The oxygenation of the earth by cyanobacteria, mentioned above, was therefore problematic for nitrogen fixing organisms. Many species retreated to suitable anoxic environments, but cyanobacteria did not have this option due to the oxygen produced by its own photosynthesis [12]. To continue to fix nitrogen, cyanobacteria evolved to either temporally or spatially separate out photosynthesis and nitrogen fixation [20, 21, 22]. For temporal separation, some cyanobacteria fix nitrogen at night and photosynthesize during the day while others, such as *Trichodesmium* spp., fix nitrogen in the morning and photosynthesize in the afternoon [23]. Some filamentous cyanobacteria were able to spatially separate photosynthesis and nitrogen fixation by specializing a fraction of the cells in the filament to fix nitrogen. These cells are known as heterocysts. Heterocyst differentiation evolved more than three billion years ago, relatively late in the evolutionary history of cyanobacteria [24].

Heterocysts are possibly the earliest example of mutualistic differentiation where a differentiated cell provides nutrients to other cells, and receives other nutrients from the other cells [25]. Heterocysts evolved in filamentous cyanobacteria and do not occur in unicellular species or unicellular mutants of filamentous species.

The study of heterocysts has used many different cyanobacterial species. The primary species used in experiments today is *Anabaena* sp. PCC 7120, which I will often shorten to PCC 7120. It has been used to study a wide range of characteristics of filamentous heterocystous cyanobacteria, from the heterocyst pattern [26] to transport [7] to gene expression analysis [27] and is chosen because it only has vegetative and heterocyst cells, and not other cell types, and because it has been developed to become a convenient genetic system. Other species that have been used are *Anabaena variabilis* [28], *Anabaena oscillarioides* [8], and *Anabaena flos-aquae* [29], among others. Much of the older work has been done with these other species. Because comparable studies have often not yet been redone with PCC 7120, we will use the results of studies with these other species to both parameterize our models and compare to our results. We do this under the assumption that the relevant data does not drastically change between these species.
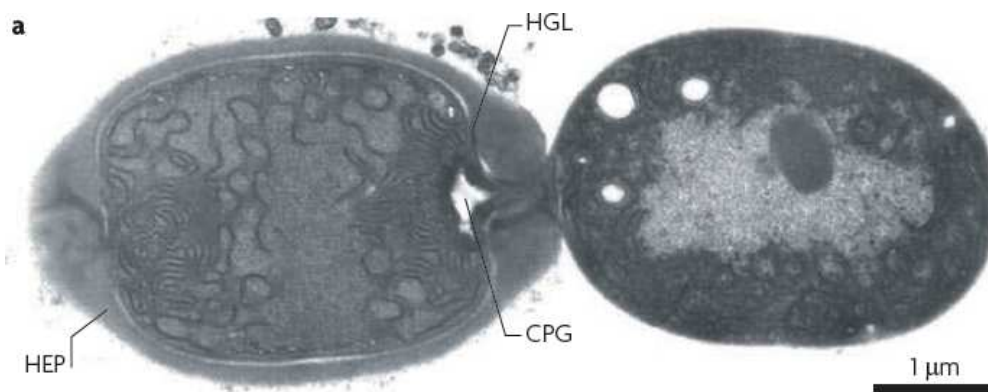
Figure 2.2: A transmission electron micrograph with a heterocyst on the left and a vegetative cell on the right from a filament of *Anabaena* PCC 7120. 'CPG' indicates the location of the heterocyst's polar cyanophycin granule. 'HEP' indicates the heterocyst polysaccharide layer, and 'HGL' indicates the heterocyst glycolipid layer. Figure from Flores and Herrero [15].

2.2   Heterocysts

Once differentiated, heterocysts no longer divide, but remain metabolically active [12]. Heterocyst death leads to filament breakage at the heterocyst [19]. The development and maintenance of heterocysts involves between 15% and 25% of the cyanobacterial genome [30]. Heterocysts have many adaptations to maintain the microoxic environment necessary for nitrogen fixation, resulting in them being darker and larger than vegetative cells. Heterocysts are surrounded by a thick envelope, composed of glycolipids and polysaccharides [19], which is impermeable to gases [15, 31]. Molecular nitrogen, N_2, is necessary for nitrogen fixation. If N_2 is able to enter the heterocyst, so is O_2, because the van der Waal's radii of nitrogen and oxygen are similar [32]. The heterocyst envelope decreases the ability of both gases to enter heterocysts to reduce the amount of O_2 that enters but still allows enough N_2 to be available for nitrogen fixation [32]. The 4 N_2:1 O_2 ratio in the atmosphere helps as well [32]. In heterocysts, photosystem II has been dismantled [33, 34], likely to avoid the oxygen produced by its water splitting activity [3, 15, 31]. Photosystem I remains, as its phosphorylation activity is important to heterocyst bioenergetics [15]. Heterocysts also have high oxidase activity to consume the O_2 that does get into heterocysts.

Nitrogenase fixes nitrogen in the microoxic environment inside the heterocyst.

That nitrogen fixation is restricted to heterocysts [35] has been demonstrated by a variety of means including ^{55}Fe labelling of nitrogenase [36] and immunoelectron microscopy [37, 38, 39, 40, 41]. Nitrogenase is encoded by three polypeptides in the *nif* operon [42].
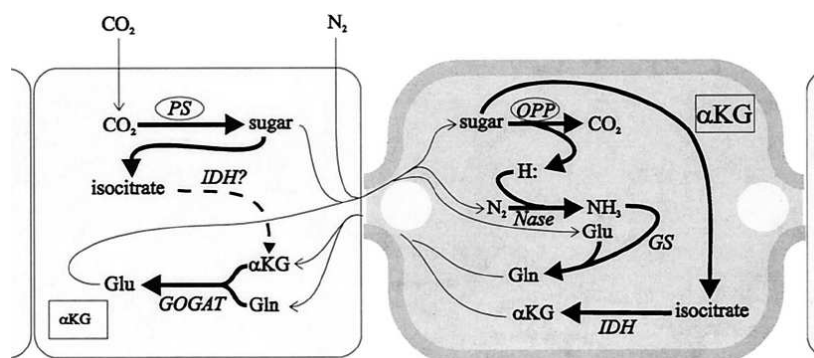
Figure 2.3: Some important metabolic interactions between heterocysts and vegetative cells. The left cell is a vegetative cell, and the right is a heterocyst. Carbon dioxide (CO_2) is fixed in vegetative cells by photosynthesis (PS); some of this fixed carbon becomes α-ketoglutarate (αKG), which is also known as 2-oxoglutarate (2-OG). α-ketoglutarate combines with glutamine (Gln) using the enzyme glutamate synthase (GOGAT) to form glutamate (Glu). In the heterocyst, sugar from the vegetative cell is used to provide reductant for the nitrogenase (Nase) and produce ammonia (NH_3) which then becomes part of glutamine. The glutamine is then the form of fixed nitrogen which moves from heterocyst to the vegetative cell. Figure from Meeks and Elhai [31].

Ammonium is the initial product of nitrogen fixation, but it is quickly incorporated into other forms of fixed nitrogen [43]. For transport back to the vegetative cells, glutamine synthetase catalyzes the formation of glutamine from ammonium and glutamate [44]. Glutamate synthase catalyzes the formation of two glutamate molecules from a molecule of glutamine and a molecule of 2-oxoglutarate (2-OG) (also known as α-ketoglutarate (α-KG)) [44]. It is thought that glutamine synthase converts ammonium into glutamine in the heterocyst while the conversion of glutamine and 2-OG to glutamate occurs primarily in vegetative cells (heterocysts are deficient in glutamate synthase) [45]. Lacking the enzyme Rubisco and photosystem II, heterocysts cannot fix CO_2 and rely on vegetative cells to supply fixed carbon to support nitrogen fixation [12]. It is thought that vegetative cells send carbon to the heterocysts in the form of sucrose [3, 15, 31]. Much of this sucrose is converted to

2-oxoglutarate (2-OG) and then has nitrogen atoms added to become glutamate and glutamine [25]. The possible physical connections for intercellular movement of fixed carbon and fixed nitrogen compounds are discussed in Section 3.1.

2.3   Heterocyst patterns
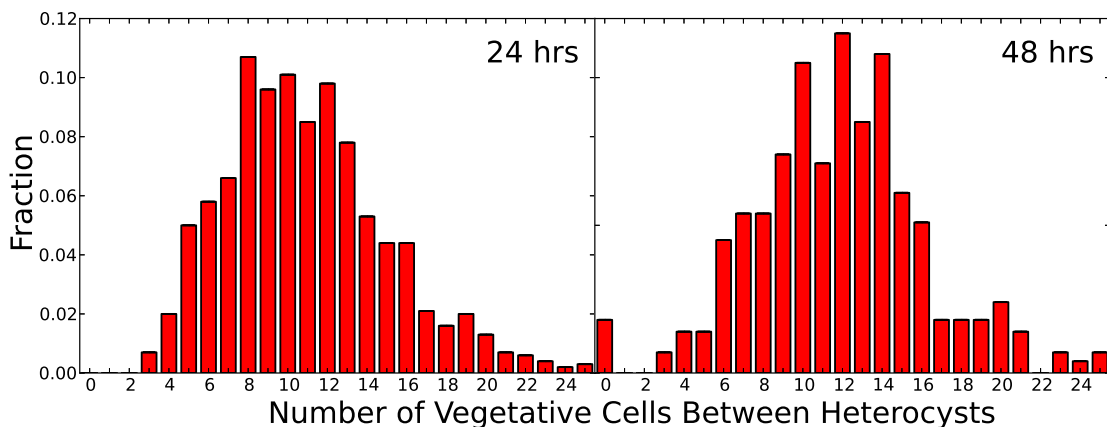
Figure 2.4: Heterocyst spacing pattern for wild type PCC 7120 24 and 48 hours after nitrogen deprivation. Data from Toyoshima *et al* [46].

As discussed in Section 1.1, we are most interested in the regularly spaced heterocyst pattern. The spacing distribution of this pattern has been measured by several groups for both wild type and knockout strains [26, 46, 47]. In wild type strains the distribution spans from roughly 5 cells to roughly 20 cells between heterocysts, with a peak at approximately 10 cells. It has been speculated that the pattern is to ensure efficient production and distribution of fixed nitrogen [12].

This thesis moves towards understanding the pattern of differentiation. This pattern is an excellent choice to study experimentally because the pattern only involves two cell types (which are visually distinguishable from one another), cell differentiation is induced by the lack of a single nutrient, and the system is one dimensional.

The heterocyst pattern is completely formed by approximately 24 hours after nitrogen deprivation. The pattern at later times (>48 hours) is quantitatively different than the pattern at 24 hours, but overall quite similar [26], as illustrated in Figure 2.4. At later times the pattern is maintained by the differentiation of new, intercalary,

heterocysts between subsequent existing heterocysts as the number of vegetative cells between them increases through division [12].

2.4 Heterocyst differentiation, the canonical experiment, and the genetic network

Heterocyst development is metabolically expensive, and heterocysts do not grow or divide, making it important that no more cells than necessary differentiate [12]. This is achieved by having flexibility in the development process using prospective heterocysts, or proheterocysts, which, before a point of no return in the differentiation process known as commitment, can halt differentiation and revert to a vegetative state [12]. Visually, proheterocysts are clearly different from vegetative cells, but lack the fully thickened walls and polar bodies of mature heterocysts [12].

The canonical experiment to investigate the development and pattern of heterocysts involves growing filaments in fixed nitrogen replete medium, and then transferring them to a medium lacking fixed nitrogen. Once in this depleted medium the development of proheterocysts and mature heterocysts occurs within 24 hours. Genes may also be knocked out and/or chemicals added to observe their effect on heterocyst development, frequency, pattern, as well as other quantities and characteristics.

Visible development of proheterocysts begins at 6-8 hours after fixed nitrogen deprivation (in a species with doubling time of 15-20 hours) [48] or 6-12 hours (in a species with doubling time of 15-25 hours) [49]. Proheterocysts mature and gain the thickened cell walls and polar bodies of mature heterocysts, with mature heterocysts reaching a maximum in frequency at 25-30 hours [48]. Because there is no fixed nitrogen being produced prior to maturity, the pattern formation, at least initially, does not depend on nitrogen fixation [50].

Most of the understanding of the process of heterocyst selection and differentiation cannot be determined by visual inspection in bright-field microscopy and uses fluorescent reporters to observe biochemical and genetic events.

In conditions of limited fixed nitrogen, 2-oxoglutarate (2-OG) levels become elevated because glutamate and glutamine are limited. The main metabolic role of 2-OG is to provide a carbon skeleton for the incorporation of nitrogen and so it links the carbon and nitrogen metabolisms [19]. The elevated 2-oxoglutarate levels which

---

The first gene important to heterocyst differentiation and not just the nitrogen deprivation response is *hetR*. *hetR* is regarded as the master switch of heterocyst differentiation [25] and knockout mutants show no sign of heterocyst differentiation [54]. Induction of *hetR* begins 2 hours after nitrogen deprivation, and by 3 hours its transcription rate is strongly enhanced [55]. Once it is expressed, *hetR* positively autoregulates, and also positively regulates *ntcA* and vice versa [19]. It has been speculated that *hetR* primarily determines whether or not a cell will complete differentiation, but does not itself direct differentiation [19].

Because *hetR* is expressed after nitrogen deprivation, positively autoregulates, and acts as the master switch controlling heterocyst differentiation, it must be turned off in the majority of cells that will not complete differentiation. The gene *patS* encodes a short peptide that acts as a diffusible negative signal of differentiation [6, 26] and is thought to inhibit *hetR* expression [49]. Indeed, strains without a functional *patS* gene have far more heterocysts after 24 hours and exhibit the multiple contiguous heterocyst (Mch) phenotype [6, 26]. Interestingly, *patS* expression is positively regulated by HetR [56]. PatS levels are observed using GFP fusions; 6 hours after nitrogen deprivation individual cells and small groups of cells show increased fluorescence, and after 8 hours many cells showed even higher fluorescence, and in a pattern similar to mature heterocysts [26]. At 10 hours most clusters of fluorescence have resolved to single cells [26].

Since PatS inhibits heterocyst differentiation, an interesting question is how *patS* does not inhibit the differentiation of the cell which is expressing it. It has been suggested that the product of *patA* expression attenuates the inhibition from PatS [51]. Transcription of *patA* is increased between 3 and 6 hours after fixed nitrogen deprivation [57]. PatA is not thought to be transferred to other cells, unlike the active peptide fragments of PatS. In this way, proheterocysts can inhibit adjoining cells from progressing to become heterocysts while not inhibiting themselves.

Cells which are undergoing differentiation into a heterocyst will at some point commit to becoming a heterocyst and no longer be able to halt differentiation and revert to a vegetative cell. To determine commitment times, either ammonia or PatS peptides are added to the growth medium of the cyanobacteria at different times after nitrogen deprivation, and then the number of heterocysts can be counted at 24 hours.

Yoon and Golden [26] found no heterocysts at 24 hours if the ammonia or PatS was added before 8 hours, and little change from experiments with no additions if the ammonia or PatS was added after 14 hours. The cells are therefore committing to differentiation between 8 and 14 hours after nitrogen deprivation. Similar experiments by Bradley and Carr [48] find commitment times between 5 and 10 hours with a faster growing species. The 8 to 14 hour commitment time is very similar to the increased fluorescence and patterning of fluorescence of GFP PatS fusions mentioned above.

Committed cells will complete differentiation into mature nitrogen-fixing heterocysts. Bradley and Carr [48] see a growth effect indicating the beginning of nitrogen fixation after 20 hours of nitrogen deprivation in *Anabaena cylindrica*. Maldener *et al* [58] observe a slight increase in nitrogen fixation after 13 hours and a very rapid increase after 22 hours in *Anabaena variabilis*. Golden *et al* [42] measured the expression of the *nif* genes (that encode the three proteins that compose nitrogenase) every six hours in PCC 7120. At 18 hours there is no signal, at 24 hours there is a weak signal, and at 30 hours there is a signal that is very strong in comparison to the signal at 24 hours. The expression of the heterocyst ferrodoxin, which directs electron flow to *Anabaena* nitrogenase from several sources, is also not transcribed in heterocysts until 18-24 hours after nitrogen deprivation in PCC 7120 [59]. It seems that in PCC 7120 nitrogen fixation begins between 18 and 24 hours and increases to much higher levels later, with faster growing strains such as *Anabaena cylindrica* fixing nitrogen even earlier.

In strains with doubling times of 15-20 hours, the fraction of cells that are proheterocysts reaches a maximum of 5-10% by 12-15 hours after nitrogen deprivation, and mature heterocysts increase in frequency until a maximum at 25-30 hours after nitrogen deprivation [48]. Strains that grow with N_2 as a nitrogen source are referred to as growing diazotrophically. The diazotrophic growth rate has been measured to be approximately 3/4 of the growth rate in nitrogen replete media [60].

Another gene, *hetN*, is important to the pattern after 24 hours of nitrogen deprivation. While *hetN* mutants exhibit a wild type heterocyst pattern after 24 hours of nitrogen deprivation, they exhibit an Mch phenotype after 48 hours [61]. HetN prevents activation of *hetR* and suppresses differentiation [62]. HetN is not present until sometime betwen 6 and 12 hours after nitrogen deprivation [63] and is thought to act

similarly to PatS but plays a role in maintaining the pattern, rather than initializing it.

To summarize the developmental process as it is presented here, it begins with nitrogen deprivation. This then leads to the elevation of 2-OG levels. These cause the expression of *ntcA* and *hetR*, which reinforce the expression of themselves and each other. HetR turns on *patS* expression, which can diffuse to other cells and inhibits *hetR* expression; local *patA* expression weakens this inhibition. The combination of HetR, PatS, and PatA determines which cells will commit to differentiation and which will revert to vegetative states. NtcA guides development in cells that commit until they are mature and express the genes to fix nitrogen. Mature heterocysts express *hetN* to maintain the heterocyst pattern.

There are many genes involved in forming the pattern and controlling heterocyst differentiation other than those we have mentioned [15, 19], but we have focused on those that are strongly supported in the literature as being important for the establishment of the pattern. Other genes are instead thought necessary for heterocyst development rather than pattern determination.

We believe that much of the variability between experiments in the timing of specific events can be attributed to the different growth rates with different growth conditions, though some may be due to the use of different species of heterocystous filamentous cyanobacteria. Quantitative data is not yet of sufficient quality or quantity, so we will use results from various conditions and species to estimate quantitative parameters with the understanding that the resulting parameters will be only approximate.

2.5   Nitrogen storage

2.5.1   Cyanophycin

Cyanophycin is a branched polypeptide of 1:1 arginine to aspartic acid that is a product of non-ribosomal synthesis (*cphA* encodes a cyanophycin synthesis enzyme, and *cphB* encodes a cyanophycin degradation enzyme [65]), and is a nitrogen-rich reserve material found in most cyanobacteria [66, 67]. Cyanophycin serves as a cellular nitrogen reserve and is mostly found in granules [67]. In nitrogen deprivation experiments
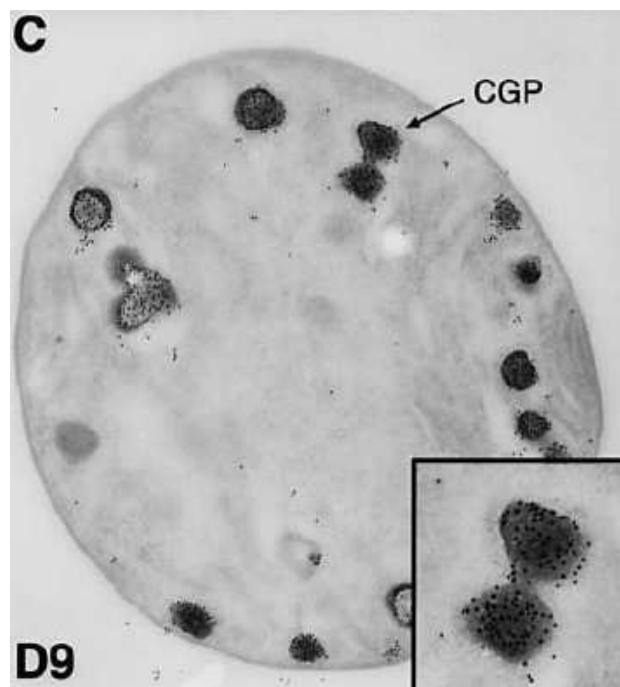
Figure 2.6: Cross section of the cyanobacteria *Cyanothece* ATCC 51142. 'CGP' indicates a cyanophycin granule. Black dots are 10nm gold labels conjugated to antibodies. Figure from Li *et al* [64].

cyanophycin is degraded before phycobiliprotein (introduced in the next subsection, 2.5.2) [67].

One experiment has found that when cyanobacteria are grown in a 12h light, 12h dark cycle, cyanophycin had accumulated to 20% of the total protein [64]. Cyanophycin has been reported as between zero and 8% of dry weight [68], as well as 568 μg of arginine per μg of chlorophyll more than cyanophycin-free species [65].

It has been suggested [69] that the uneven accumulation of cyanophycin would cause the cells with the lowest supply to notice starvation first, but a reason for uneven accumulation was not provided.

2.5.2 Phycobiliprotein

Phycobiliproteins are the most abundant proteins in the cyanobacterial cell, and are typically found in phycobilisomes [67]. Phycobilisomes are composed of light harvesting phycobiliprotein pigments that absorb energy to transfer to photosystem II [67].

Nitrogen starvation results in the activation or new synthesis of proteases that break down phycobiliproteins [69]. During the first hours after fixed nitrogen deprivation, phycobiliprotein degradation appears to start in all vegetative cells but only completes in the heterocysts, where the phycobiliprotein content goes to nearly zero [70]. Mature heterocysts seem to regain phycobiliproteins later on [70]. The phycobiliprotein in vegetative cells degrade only transiently and partially.

Cellular levels of phycobiliprotein markedly decrease when cells are deprived of a source of combined nitrogen [67]. Cells depleted of phycobiliproteins can continue to photosynthesize because sufficient chlorophyll a remains with photosystem II [67]. Phycobiliproteins therefore represent an additional potential pool of amino acids to be utilized with no irretrievable damage for a cell starved of fixed nitrogen.

2.6 Modeling biological pattern formation

Chemical and biological pattern formation studies often refer to Turing [71], who showed that under certain conditions two interacting chemicals can generate a stable inhomogeneous pattern.

Theoretical models for biological pattern formation must give satisfactory answers to two questions: 1) How can a system give rise to and maintain large-scale inhomogeneities like gradients even when starting from more or less homogeneous conditions?; and 2) How do cells measure the local concentration in order to interpret their position in a gradient and choose the corresponding developmental pathway [5]?

Both of these questions apply to heterocyst differentiation in filamentous cyanobacteria. Although it is known which genes must have their levels rise to become a heterocyst, it is not known how the filament decides which cells are destined to become heterocysts and increase the expression of these genes. It has been suggested that the cell cycle [72, 73] or differences in nitrogen storage [69] could contribute to the breaking of this symmetry. With regard to the second question, it is also unknown how all the relevant proteins interact with one another, in particular, how PatS does not inhibit the differentiation of the cell which produced it.

Generally, there are two features important to biological pattern formation: local self-enhancement and long-range inhibition. Self enhancement is essential for local inhomogeneities to be amplified. A substance is said to be self-enhancing if a small
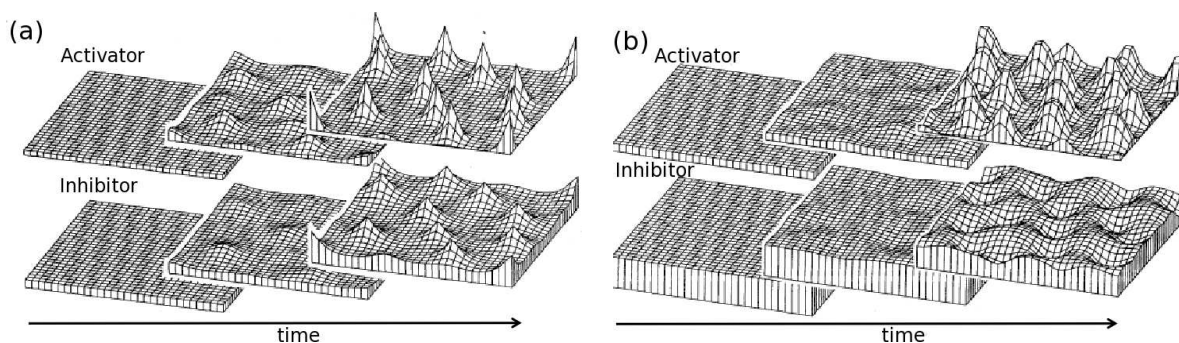
Figure 2.7: The progression of inhomogeneities in a model where (a) the inhibitor is produced by the activator, and (b) the inhibiting substrate is consumed by the activator. In both (a) and (b) the top sequence is the activator distribution and the bottom is the inhibitor distribution, with initial, intermediate, and final patterns going from left to right. Figure from Koch and Meinhardt [5].

increase over the steady state concentration induces further increase. Self enhancement must be accompanied by a fast diffusing inhibitor to avoid overall activation. There are two options for an inhibitory substance: one which is produced by the activator, the other is a substrate consumed by the activator. For the correct range of parameters, inhibitors produced by the activator tend to produce sharp peaks of activator while substrates which are consumed tend to have rounded, gradual activator maxima. Insertion of new maxima in systems where the activator produces the inhibitor occur where the inhibitor is too low to suppress the activator production [5].

Both of these paradigms fit heterocyst differentiation well. During initial differentiation, fixed nitrogen, which inhibits heterocyst differentiation, is consumed, leaving clusters of proheterocysts. This fits as a substrate (fixed nitrogen) consumed by the activator (starving cells). For the other type, the fast diffusing inhibitor in the system is PatS, and it is produced by the activator, HetR. In wild type filaments, there are no Mch heterocysts, indicating the sharp peak consistent with the activator produced inhibitor.

2.7 Previous heterocyst pattern modeling efforts

There have been several modeling studies of fixed nitrogen and heterocyst patterns in cyanobacteria in the past.

In 1974 Wolk *et al* [11] developed a quantitative model of fixed nitrogen transport and incorporation along the cyanobacterial filament, including diffusive transport and fixed nitrogen consumption in vegetative cells. The model treated the filament as a continuous medium, treated the filament as a single compartment, did not include stochasticity or growth, and assumed that consumption was linearly related to the fixed nitrogen concentration at all concentrations.

In 1975, Wolk and Quine [74] simulated random heterocyst placement in a filament with an inhibition zone and found that the resulting heterocyst pattern was qualitatively consistent with the experimental heterocyst pattern. This identified lateral inhibition as a plausible mechanism behind the observed heterocyst patterns. The model did not include fixed nitrogen or growth.

In 2002, Meeks and Elhai [31] found that for random placement the heterocyst spacing distribution is maximum for neighbouring heterocysts (spacing of zero) and monotonically decreases for larger spacings. This demonstrated that random spacing is in clear disagreement with the experimentally measured wild type distribution and also does not correspond to the distributions of any observed mutants.

In 2007, Allard *et al* [56] developed a stochastic cellular computational model of the cyanobacterial filament with growth, fixed nitrogen dynamics, and heterocyst selection and differentiation. The work found the numerical range of parameters for their model that most closely reproduced the experimental heterocyst spacing distribution for a particular mutant cyanobacteria. The model only assumed periplasmic transport (no cytoplasm to cytoplasm transport) and did not investigate the possibility of local fixed nitrogen starvation.

Our approach is most similar to that of Allard *et al* [56]. Our cyanobacteria model has growth, fixed nitrogen dynamics, heterocyst selection and differentiation. In contrast to Allard *et al*, we consider both cytoplasmic and periplasmic transport and we do not find our model parameters as an outcome of our work, but instead our models use parameters that have been estimated from the literature. We investigate transport and consumption, as Wolk *et al* did in 1974 [11], and spacings, as Wolk and Quine did in 1975 [74], but in the context of our more fleshed out model. We also add other elements to the model that have not been considered before, such as storage. Our choice of model is guided by what we want to investigate. To

investigate the nitrogen distribution and transport, the model needed to have fixed nitrogen transport and consumption, growth, and a model for existing heterocysts. To explain the advantage of the heterocyst pattern we of course needed to add a model for the dynamic placement of heterocysts. Finally, to consider the formation of the heterocyst pattern upon fixed nitrogen starvation, we further developed the heterocyst placement model with lateral inhibition and use fixed nitrogen storage to break the symmetry.

Chapter 3

Nitrogen Distribution in the Cyanobacterial Filament

3.1   Motivation and background

This chapter is based on a paper we have published in Physical Biology, volume 9, page 016007, 2012 [75]. IOP Publishing has given me permission to reproduce the contents of the paper in this thesis (see Appendix B). For this paper my contributions were to write the simulation, do the analytical and computational work, generate the figures, and write the first draft. I was also an equal partner with my supervisor in developing the model and revising the paper.

The first thing we wanted to understand was the dynamics of fixed nitrogen, the limiting nutrient for heterocyst differentiation, in the filament. To help understand this we looked in the literature for experimental nitrogen distributions in the filament and found two: the distribution of Wolk *et al* [11], which I will refer to as the Wolk data, and the distribution of Popa *et al* [8], which I will refer to as the Popa data. An immediate challenge was that these two distributions are quite different.

The study of Wolk *et al* in 1974 [11] provided growing filaments of *Anabaena cylindrica* with dinitrogen gas composed of ^{13}N (half-life of less than 10 min) for 2 min before being fixed and imaged. Radioactive tracks from decaying ^{13}N were used to approximate the distribution of labelled fixed nitrogen in the filament. Wolk found that the fixed nitrogen was peaked at the location of the heterocysts - consistent with diffusive transport away from a heterocystous source.

The more recent study by Popa *et al* in 2007 [8] provided growing filaments of *Anabaena oscillarioides* with dinitrogen gas composed of (stable) ^{15}N. They used NanoSIMS (nanometer-scale secondary-ion mass spectrometry) to obtain the distribution of ^{15}N in the filament. After 4 hours of growth, they found *dips* in the nitrogen distribution at the locations of the heterocysts, and a noisy plateau between them. There was no evidence of a concentration gradient consistent with diffusive transport.

The clear contrast between these two distributions is that Wolk found a peak

in the nitrogen distribution at the heterocysts, as is expected for the source, while Popa found a dip. It is not clear how to reconcile these two distributions. A difference between the two experiments is that Wolk used cyanobacterial species *Anabaena cylindrica*, while Popa used *Anabaena oscillarioides*. Neither of these two species is the more commonly used model organism *Anabaena* sp. PCC 7120. It is tempting, but not necessary, to attribute the qualitative differences between the Wolk and Popa distributions to distinct species-species physiology. We expect that fixed nitrogen transport is qualitatively similar in all heterocystous filamentous cyanobacterial species and we investigated the quantitative fixed nitrogen patterns expected from the two experimental approaches by using a single model.

To understand the distribution of fixed nitrogen in the cyanobacterial filament we must model several processes, one of which is the transport of fixed nitrogen between the cells in the filament. Transport can either be from cytoplasm to cytoplasm, or periplasm to periplasm. There is evidence supporting either process in the literature.

Direct cytoplasmic connections between adjacent vegetative cells were first observed in electron microscopy images of small (about $50\overset{\circ}{A}$ in diameter) holes in the septal cell membrane called microsplasmodesmata, after the much larger plasmodesmata of plant cells [76, 77, 78]. Large pores have also been observed connecting heterocysts to adjoining vegetative cells [79]. More recent measurements of direct cytoplasmic transport have been made using the small exogenous fluorophore calcein (molecular mass 623 Da) in *Anabaena* sp. PCC 7120 using the technique of fluorescence recovery after photobleaching (FRAP) [7]. The FRAP technique photobleaches a specific region, whether smaller than a cell or encompassing many cells, and then allows the fluorescence to recover. Any fluorescence recovery is due to the diffusion of fluorophores that were not in the photobleached region at the time of photobleaching, and can often allow the quantification of transport. Mullineaux *et al* [7] photobleached individual cells whose fluorescence then recovered on the timescale of tens of seconds and allowed the constant for a transport model between cells to be estimated. Since calcein is not native to cyanobacteria, specific transport mechanisms are unlikely, and any passive diffusive transport mechanism could also be used to transport fixed nitrogen products. The protein SepJ (also known as FraG) appears to be important for direct cytoplasmic transport. SepJ has a coiled-coil domain that prevents filament

fragmentation and a permease domain that is required for diazotrophic growth [80]. SepJ also appears to be needed for microplasmodesmata formation [81]. There is no evidence that the operation of SepJ requires ATP or has any energy requirements for action, meaning that the direct cytoplasm to cytoplasm transport of fixed nitrogen and other small molecules is likely to be by passive diffusion.

Electron microscopy has also shown that the periplasmic space of cyanobacterial filaments is continuous, that is, the outer cell membrane does not come between the inner cell membranes of adjacent cells and the outer cell membrane is common to all cells in the filament [82]. Supporting these observations, the transport of GFP (molecular mass of about 27kDa) from one periplasm to the periplasm of an adjacent cell has been observed in FRAP experiments [83], although it is controversial [76, 84, 85]. The outer membrane has been shown to provide a permeability barrier to fixed nitrogen compounds [86]. Also, knocking out amino acid permeases (transporters) between the periplasm and cytoplasm has led to reduced diazotrophic growth [87, 88, 60]. The continuous periplasm, GFP transport, and amino acid permease relation to growth all suggest that some amount of fixed nitrogen is transported through the periplasmic space.

There is evidence for both cytoplasmic and periplasmic fixed nitrogen transport and no measurements of the relative fluxes. The evidence for cytoplasmic transport shows rapid transport between cells and is consistent with diffusion. There is likely some fixed nitrogen in the periplasm, and lateral gradients would lead to diffusion. There is no evidence that the gradients in the periplasm are significant or that the diffusion is rapid.

Previous modeling efforts relating to this work are reviewed in Section 2.7.

3.2   Model

3.2.1   Fixed nitrogen transport

Our fN transport and incorporation model is similar to that of Allard *et al* [56] and tracks the total amount of cytoplasmic fixed nitrogen, $N_C(i,t)$, in each cell i versus time t:

$$\frac{d}{dt}N_C(i,t) = \Phi_C(i) + D_I N_P(i,t) - D_E N_C(i,t) + G_i \qquad (3.1)$$

We similarly track the periplasmic fixed nitrogen, $N_P(i,t)$:

$$\frac{d}{dt}N_P(i,t) = \Phi_P(i) - (D_I + D_L)N_P(i,t) + D_E N_C(i,t) \tag{3.2}$$

Φ_C and Φ_P are the terms for the net cytoplasmic and periplasmic flux respectively, determining transport with the neighbouring cells and are described in more detail below. Φ_C is proportional to the constant D_C and Φ_P is proportional to the constant D_P. D_I is the coefficient for import from the periplasm to the cytoplasm, D_E is the coefficient for export from the cytoplasm to the periplasm, and D_L is the coefficient for losses from the periplasm to the external medium. These transport processes are shown schematically in Figure 3.1. Note that any fixed nitrogen transferred between the cytoplasm and periplasm, with D_I or D_E, shows up with opposite sign in Equations 3.1 and 3.2 so that the total amount of free fixed nitrogen is unchanged by the exchange. In addition to these transport terms, there is a source/sink term G_i in Equation 3.1 that describes fixed nitrogen production and consumption in the cytoplasms of heterocysts and vegetative cells, respectively. This G term is discussed more extensively in Section 3.2.2.
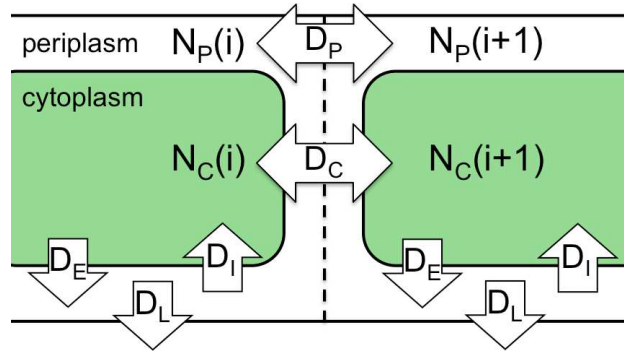
Figure 3.1: Schematic of fixed nitrogen dynamics as represented by Equations 3.1 and 3.2. Fixed nitrogen amounts are indexed to each cell by i, with $N_C(i)$ the cytoplasmic fixed nitrogen and $N_P(i)$ the periplasmic fixed nitrogen of each cell. Each shaded green region represents the cytoplasms of two adjacent cells. D_C describes the magnitude of the cytoplasm to cytoplasm transport, D_P the periplasm to periplasm transport, D_I the import from the periplasm to the cytoplasm, D_E the export from the cytoplasm to the periplasm, and D_L the loss from the periplasm to outside the filament.

Diffusive transport between adjacent cytoplasms and periplasms is described by the flux terms Φ. As in the models of Allard *et al* [56] and Mullineaux *et al* [7], we

assume every compartment in our model is well mixed and that transport is limited by the junctions between cells. Supporting this, there is no evidence in the images of fixed nitrogen by Popa *et al* [8] or of calcein by Mullineaux *et al* [7] of subcellular gradients. Fick's law states that the net diffusive flux between two cells will be proportional to the density difference between the cells. If a cell's neighbours were empty of fixed nitrogen, it would have two outgoing fluxes, one to the left neighbour, $\Phi_{L,out}$, and one to the right neighbour, $\Phi_{R,out}$. If a cell were empty of fixed nitrogen, it would have two incoming fluxes, one from the left neighbour, $\Phi_{L,in}$, and one from the right neighbour, $\Phi_{R,in}$. The actual net flux into a cell is the sum of these four fluxes:

$$\Phi = \Phi_{L,in} - \Phi_{L,out} + \Phi_{R,in} - \Phi_{R,out}$$

The flux is proportional to the concentration of fixed nitrogen. With the cross section of the cells in the filament remaining constant as they grow and divide, we use the one dimensional concentration $N(i,t)/l(i,t)$ (l is the cell length) and absorb the cross section into the proportionality constant. The cytoplasmic flux in our model is proportional to D_C.

$$\Phi_C = D_C \left(\frac{N_C(i-1,t)}{l(i-1,t)} - \frac{N_C(i,t)}{l(i,t)} + \frac{N_C(i+1,t)}{l(i+1,t)} - \frac{N_C(i,t)}{l(i,t)} \right)$$

$$= D_C \left(\frac{N_C(i-1,t)}{l(i-1,t)} + \frac{N_C(i+1,t)}{l(i+1,t)} - 2\frac{N_C(i,t)}{l(i,t)} \right)$$

Similarly for the periplasmic flux.

$$\Phi_P = D_P \left(\frac{N_P(i-1,t)}{l(i-1,t)} + \frac{N_P(i+1,t)}{l(i+1,t)} - 2\frac{N_P(i,t)}{l(i,t)} \right)$$

As long as the transport coefficients (D_C and D_P) are constant along the filament, these flux terms look like a discrete Laplacian $\nabla_d^2 \equiv N(i-1,t)/l(i-1,t) + N(i+1,t)l(i+1,t) - 2N(i,t)l(i,t)$ and lead to the diffusion equation when coarse grained (see Equation 3.5 below).

The magnitudes of D_C and D_P will be proportional to the microscopic diffusivity of fixed nitrogen but, given that we are assuming that transport is limited by the junctions between cells, is also dependent on the nature of the junctions between cells. Because of this we have simply taken experimentally measured transport constants for other molecules and scaled them to glutamine, which is the molecule that carries

fixed nitrogen from heterocysts to vegetative cells. In doing this we have assumed that glutamine and the other molecules are treated the same by the cell connections and that the transport will only change as it would for regular diffusion. By this we mean that we are assuming a Stokes-Einstein diffusivity for the diffusion, so the diffusion is inversely proportional to radius, corresponding to the inverse cube root of the molecular mass.

To estimate D_C, we will use the cytoplasm to cytoplasm transport coefficients for calcein in *Anabaena* sp. PCC 7120 filaments [7]. We choose to use the transport constants in filaments that have been starved of fixed nitrogen for 96 h, because the experiments of Wolk *et al* [11] and Popa *et al* [8] use starved filaments. Multiplying the exchange coefficient E of units s^{-1} by a cell length of 3.38 μm gives the constant in the required units of μm s^{-1}. They are then scaled from calcein to glutamine by a factor of the cube root of the ratio of the molecular masses. We obtain $D_C = 1.54$ μm·s^{-1} between two vegetative cells and $D_C = 0.19$ μm·s^{-1} between a vegetative cell and a heterocyst. We do not change these values in this thesis.

To estimate D_P, we do not have a measured value in the literature for periplasm to periplasm transport of a small molecule in *Anabaena* sp. PCC 7120. Instead, we will use the periplasm to periplasm transport of GFP in *Anabaena* sp. PCC 7120 [83] to obtain a floor on D_P and the intraperiplasm transport of GFP in *Anabaena* sp. PCC 7120 [85] to obtain a ceiling on the value of D_P.

Mariscal *et al* observed a FRAP timescale of approximately 150s for GFP moving between adjacent cells [83]. This corresponds to a D_P for GFP of approximately $L/150\text{s} \approx 0.02$ μm·s^{-1} using a cell length $L \approx 3$ μm. Scaling the transport of GFP (27 kDa) to glutamine (146 kDa) by the cube root of the molecular masses [89], i.e. by a factor of $(27 \text{ kDa}/146 \text{ Da})^{1/3} \approx 5.7$, and taking this as a lower bound, then gives $D_P \gtrsim 0.1$ μm·s^{-1}. FRAP experiments with GFP in *Anabaena* sp. PCC 7120 showed recovery in part of a single cell's periplasm after around 60 s [85] and after around 5 s [83]. This faster time of 5 s is similar to the 4 s recovery time observed in the periplasm of *E. coli* [90], which also measured an associated diffusivity of $D_{P,Ecoli} \approx 2.6$ μm^2·s^{-1}. Scaling this diffusivity to glutamine, we obtain $D_P = D_{P,Ecoli} \times 5.7/3.8\mu$m$\approx 4.4\mu$m s^{-1}. Given that glutamine has a much smaller mass than GFP, it is unlikely that it will diffuse any slower from periplasm to periplasm than when scaled from GFP. The

broad range we consider for periplasmic transport is then 0.1 μm s$^{-1} \lesssim D_P \lesssim 10\mu$m s^{-1}.

When we consider periplasmic transport, we must also consider exchange between the cytoplasm and periplasm, as well as loss from the periplasm to the external medium. We expect that uptake of fixed nitrogen from the periplasmic to the cytoplasmic compartment, controlled by D_I in Equations 3.1 and 3.2, is an active process, controlled by ABC transporters [60] or a similar system. Nevertheless, we expect that the rate of import is proportional to the periplasmic fixed nitrogen density, N_P/l, and to the number of transporters, which will themselves be proportional to the cell length l. Similarly, transport of fixed nitrogen from the cytoplasmic to the periplasmic compartment, controlled by D_E, will be proportional to the cytoplasmic density N_C/l but also to the amount of membrane or (possibly leaky) ABC transporters [91], which will be proportional to l. The result is that the transport terms with coefficients D_I and D_E are independent of the cell length l and the coefficients have units of s^{-1}.

To estimate the value of D_I (from Equations 3.1 and 3.2), we begin with the measurement of Pernil et al [60] that 1.65 nmol/(mg protein) is imported into a PCC 7120 filament in 10 min from a medium containing 1 μM glutamine. From our discussion in Section 3.2.2, the dry mass of a PCC 7120 cell is 4.89×10^{-12}g and the mass of protein, approximated at 55% of the dry mass of the cell [92], is 2.69×10^{-12}g. This implies that there is 4.44×10^{-9} nmol of glutamine, or $N_G = 2.67\times10^6$ glutamine molecules, imported in 10 min into each cell. The term in Equation 3.1 describing import from the periplasm to the cytoplasm is $\partial N_C/\partial t = D_I N_P$. If we assume that the periplasmic concentration is equal to the glutamine concentration in the external medium, then N_P can be replaced by the product of the glutamine concentration in the periplasm, $\rho_{P,G}$, the cross-sectional area of the periplasm A (for a 0.1 μm thick periplasm surrounding a cytoplasm 1 μm in radius $A = 0.21\pi$ μm^2) and the cell length l giving $N_G = D_I \rho_{P,G} l A \tau$, with $\tau = 10$ min. Solving this for D_I using $l = l_{min} = 2.25$ μm yields $D_I = 4.98$ s^{-1}. ABC transporters, known to transport amino acids [88], are asymmetric, and we assume that export is ten times weaker than import, and that D_E = 0.498 s^{-1}. It is also known that the outer cell membrane provides a permeability barrier [86], implying that the periplasmic glutamine concentration may not be as high as the external glutamine concentration, and so we also consider a D_E and D_I

pair that is ten times larger with $D_I = 49.8$ s^{-1} and $D_E = 4.98$ s^{-1}.

There is evidence that fixed nitrogen is lost from the filament to the external medium [93, 94] and we include a loss term for the periplasm with the coefficient D_L in Equation 3.2. We assume the loss rate is very small compared to the other membrane fluxes, using 1% of the lower D_I value, i.e. $D_L = 0.0498$ s^{-1}. This qualitatively corresponds to the observation of a permeability barrier in the outer membrane [86]. Our results are qualitatively unchanged in the range of $0 \leq D_L/D_I \lesssim 0.1$.

3.2.2   Cell growth and division

The cells in *Anabaena* sp. PCC 7120 have a minimum size of $l_{min} = 2.25$ μm and a maximum size of $l_{max} = 2l_{min} = 4.5$ μm [15, 84]. When the length of a cell reaches l_{max}, it is divided into two cells. Each of these two daughter cells is assigned a new random growth rate and is given half of the fixed nitrogen of the parent cell. In our investigation we found that stochastic effects were not always necessary for our conclusions and so we have both deterministic and stochastic results. For the deterministic results all cells begin with the same length, $l_0 = 2.8$ μm and have the same doubling time of $T_D = 20$ h [88, 95]. Doubling time determines the optimal growth rate by having the optimal growth rate of an individual cell, $R^{opt} = l_{min}/T_D$. For stochastic results, shown in Figures 3.3 (b) and (c), all initial cells are assigned a length between l_{min} and l_{max} according to an analytical steady state distribution of cell lengths [96]. An optimal growth rate is chosen randomly and uniformly in the range between the minimum optimal growth rate $R^{opt}_{min} = l_{min}/T_{max}$ and the maximum growth rate $R^{opt}_{max} = l_{min}/T_{min}$. The maximum cell period is defined as $T_{max} = T_D + \Delta$ and the minimum cell period is defined as $T_{min} = T_D - \Delta$. The standard deviation of the growth rate is then $\sigma_R = \sqrt{2/3}\Delta l_{min}/[(T_D + \Delta)(T_D - \Delta)]$. We expect that the coefficient of variation of the growth rate σ_R/R_{avg} is of similar magnitude to that seen in a study of mutant *Anabaena* [56], which was $\sigma_R/R_{avg} \approx 0.165$.

In Equation 3.1 there is a source/sink term G_i. For vegetative cells $G_i = G_{veg}$ is a sink term that takes into account the fixed nitrogen consumption that contributes to growth. The fixed nitrogen consumed by a cell depends on the growth rate of the

cell, and also on the amount of local cytoplasmic fixed nitrogen available for growth.

$$G_{veg} = -gR(R_i^{opt}, N_C(i,t)),\qquad\qquad (3.3)$$

g is the amount of fixed nitrogen that a cell needs to grow one μm; here we will estimate this quantity. Dunn and Wolk [97] measure a dry cell mass for one cell of *A. cylindrica* of 1.65×10^{-11} g. Cobb *et al* [98] measured the nitrogen content of *A. cylindrica* to be 5-10% of dry mass (see also [99]). We will use the typical value of 10% and scale it by volume from *A. cylindrica* to PCC 7120. *A. cylindrica* is approximately 1.5 times as wide and 1.5 times as long as PCC 7120 [15], giving $\approx 2.07\times10^{10}$ N atoms per average cell of PCC 7120. According to Powell's steady state length distribution of growing bacteria [96], the average cell size is 1.44 times as long as the smallest cell, so that $\approx 1.4\times10^{10}$ N atoms are needed for a newly born cell to double in length. This implies that $g = 1.4 \times 10^{10}/l_{min} \simeq 6.2 \times 10^9\ \mu\text{m}^{-1}$.

In our growth model we assume that individual cells grow at their optimal growth rate R_i^{opt} as long as the cell has greater than zero cytoplasmic fixed nitrogen (this assumption is discussed in Section 3.5). When at zero cytoplasmic fixed nitrogen, a cell can only grow using the flux of fixed nitrogen into the cell from neighbouring cells and the periplasm.

$$R = \begin{cases} R_i^{opt} & \text{if } N_C(i,t) > 0 \\ min\left(\Phi_{in}/g, R_i^{opt}\right) & \text{if } N_C(i,t) = 0 \end{cases}\qquad (3.4)$$

It is important to note that cells with $N_C = 0$ can still grow, but this growth is limited to what can be supported by the incoming fixed nitrogen flux: $gR = \Phi_{in} = \Phi_{L,in}+\Phi_{R,in}+D_I N_P$. A cell transitioning from $N_C = 0$ to $N_C > 0$ will have $\Phi_{in} > R_i^{opt}$; in this instance the $R_i = R_i^{opt}$.

In Equation 3.1, for heterocyst cells $G_i = G_{het}$. The heterocyst fixed nitrogen production rate G_{het} is chosen to supply the growth of approximately 20 vegetative cells and is $3.15\times10^6\ s^{-1}$ unless otherwise stated.

3.2.3   Some numerical details

Filaments were initiated with two heterocysts, separated on either side by vegetative cells, on a periodic loop. Periodic boundary conditions were used to minimize end

effects. We have presented our data with a fixed number of vegetative cells between heterocysts because of cell division. Most of our data has 20 vegetative cells between the heterocyts. This is approximately twice the typical heterocyst separation [15, 84], which allows us to investigate fixed nitrogen depletion at the midpoint, which corresponds to the location of an intercalating heterocyst.

We distinguish free from incorporated fixed nitrogen. Free fixed nitrogen is freely diffusing in the cytoplasm and can be transported to other cells, and is simply given by $N_C(i,t)$. We report this as a linear density $\rho_F \equiv N_C(i)/l_i$. Incorporated fixed nitrogen is the fixed nitrogen that has been incorporated by cellular growth through the growth term G_{veg}. This nitrogen is locked to a specific cell and its daughters. We report the incorporated concentration $\rho_I \equiv \int G_{veg} dt / l_i$, where the growth is integrated over a fixed duration. During cell division, half of the parent cell's incorporated fixed nitrogen is assigned to each daughter cell. We are interested in isotopically labeled fixed nitrogen [8, 11], and so we set all $N_C(i) = 0$ at the start of our data gathering ($t = 0$), consistent with the introduction of isotopically labeled dinitrogen to the filament that is then fixed by the heterocysts and subsequently supplied to the filament as free fixed nitrogen via G_{het}. The total fixed nitrogen, for e.g. Figures 3.4(a) and (c), in a cell is $\rho_T \equiv \rho_F + N_P/L + \rho_I$, where we include periplasmic contributions as well. The two experiments this work is in the context of, from Wolk *et al* [11] and Popa *et al* [8], could not distinguish between freely diffusing and incorporated nitrogen, and recorded only their sum. Therefore the total fixed nitrogen is the quantity that is important to compare to experimental results.

3.3 Cytoplasmic transport

3.3.1 Short time fixed nitrogen distributions

First we investigated systems with only cytoplasmic transport (i.e. $D_I = D_E = D_P = 0$) over short periods of time to be consistent with the experiments of Wolk *et al* [11]. The rate of cytoplasmic transport was held constant, with the coefficient $D_C = 1.54$ μm s^{-1} between vegetative cells, and $D_C = 0.19$ μm s^{-1} between vegetative cells and heterocysts, as described in Section 3.2.1. In Figure 3.2 we show the fixed nitrogen distributions in the filament. A heterocyst is at the left end of the distribution,
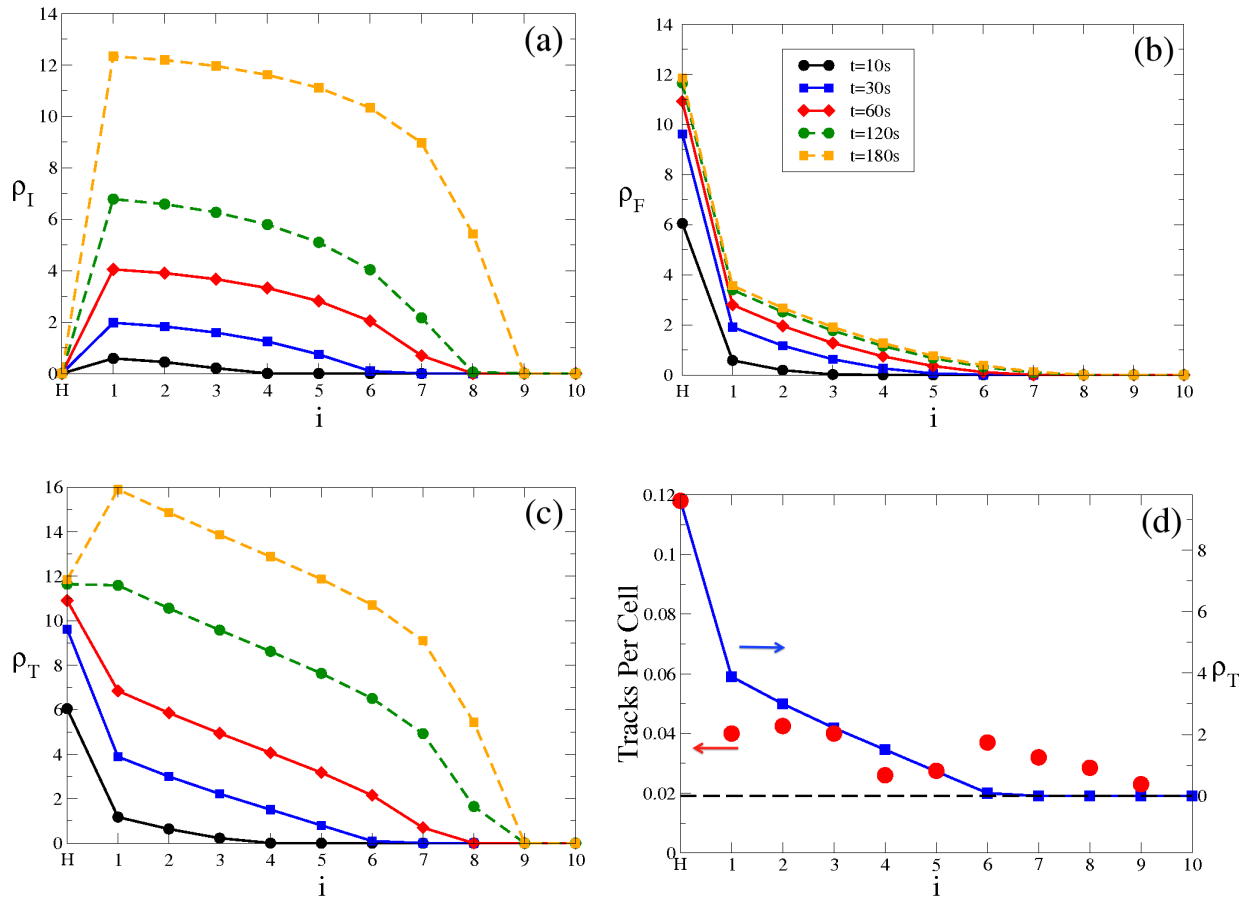
Figure 3.2: Model fixed nitrogen concentration vs. cell index i (the number of cells from a heterocyst at $i = 0$), for short intervals of labelled nitrogen fixation ranging from $10s$ to $180s$ as indicated by the legend in (b). The incorporated fixed nitrogen ρ_I is shown in (a), the free fixed nitrogen ρ_F in (b), and the total fixed nitrogen ρ_T in (c). In (d) the red circles indicate the experimental data digitized from Wolk et al [11], and the blue squares are the $t = 30s$ ρ_T data from (c) (using the right axis, scaled and shifted to best agree with experimental data). Our model data is from our deterministic model with no periplasmic transport, i.e. $D_C = 1.54 \ \mu\text{m s}^{-1}$ and $D_I = D_E = D_P = 0$.

indicated by an "H", and the vegetative cells are indexed by i, counting from the heterocyst. We varied the amount of time between the start of fixation of labelled nitrogen and the time the distribution was sampled ($t = 10, 30, 60, 120$, and 180 s, refer to the legend in the top right figure). (a) shows that the incorporated fixed nitrogen, ρ_I, increases steadily in all cells reached by the free fixed nitrogen, ρ_F, except the heterocyst. The heterocyst is not incorporating any fixed nitrogen because it is not growing. (b) shows the free fixed nitrogen, ρ_F, as it slowly extends out from the heterocyst and increases. ρ_F is non-zero seven cells away from the heterocyst, while the eighth cell receives some fixed nitrogen it does not receive enough to exceed the amount that could be used for growth. The fixed nitrogen in the heterocyst and cells one through seven increases until it reaches a steady state: at this time the rate that nitrogen is fixed by the heterocyst is equal to the rate it is used by the cells to grow. In (c) we show the total fixed nitrogen, $\rho_T = \rho_I + \rho_F$. The form of ρ_T changes in time. At the shortest times, there is a spike in ρ_T at the heterocyst and a gradient moving away; this reflects ρ_F and ρ_T is dominated by ρ_F at the shortest times. At time increases, the spike at the heterocyst flips to a dip, and the magnitude of the gradient becomes small compared to the magnitude of the total fixed nitrogen; this reflects ρ_I and ρ_T has more of the character of ρ_I as the time increases.

The radiographic technique used by Wolk *et al* [11] is unable to distinguish between free and incorporated fixed nitrogen, and can only measure the total. Therefore it is appropriate to compare the ρ_T distribution of Figure 3.2(c) to the experimental data. We have shown the average of two sets of experimental data in Figure 3.2 (d) as filled red circles (left scale). We have also included a qualitatively similar ρ_T curve from $t = 30$ s as the blue curve (right scale). We have adjusted the ρ_T scale so that the two curves agree at the location of the heterocyst and with the experimental background far from the heterocyst. We have reproduced the main feature of the Wolk data for short exposure to fixed nitrogen, the spike at the heterocyst, and the other scales are also approximately correct. Wolk's data was taken after 120 s of exposure to labelled fixed nitrogen, but our curve that best matches this data is from after 30 s. This is because in our model, heterocysts immediately provide fixed nitrogen to be transported, but this is not the case for real heterocysts. Glutamine levels in *Anabaena cylindrica* linearly increase during the first minute of exposure to labelled

dinitrogen and other fixed nitrogen products appear after 2 min [43]. This 1-2 min lag between exposure to dinitrogen and its transport and incorporation is consistent with the 90 s difference between the curves in Figure 3.2(d).

3.3.2 Long time fixed nitrogen distributions

We again examined systems with only cytoplasmic transport, but at long times comparable to those of Popa *et al* [8].

In the last section we found that after approximately 120 s, the free fixed nitrogen distribution had reached a steady state. To find an analytic expression for this free fixed nitrogen distribution we approximate the filament as a continuous diffusive medium rather than discrete cells. The diffusion and consumption of free fixed nitrogen in one dimension is

$$\partial \rho_F / \partial t = D \nabla^2 \rho_F - c \tag{3.5}$$

where ρ_F is the one dimensional concentration of free fixed nitrogen, D is the diffusivity (which is related to our cytoplasmic transport parameter D_C) of free fixed nitrogen and c is the consumption of free fixed nitrogen per unit length (related to gR from our growth model). For the steady state ($\partial \rho_F / \partial t = 0$) the general solution is quadratic against the distance x. We define $x = 0$ as the location of a heterocyst. When in steady state, the flux out of the heterocyst in each direction is equal to half the fixed nitrogen production of the heterocyst. Combining with Fick's Law, we have $G_{het}/2 = -D \partial \rho / \partial x$. The flux out of the heterocyst must also equal the total fixed nitrogen consumed by the vegetative cells on either side of it. If x_0 is the distance from the heterocyst that the fixed nitrogen reaches and where growth stops, then $c x_0 = G_{het}/2$ and $\rho_F(x_0 = G_{het}/(2c)) = 0$. Applying these two boundary conditions to the general quadratic solution we get

$$\rho_F(x) = \frac{1}{2} \frac{c}{D} x^2 - \frac{G_{het}}{2D} x + \frac{G_{het}^2}{8cD}, \tag{3.6}$$

for $x \leq x_0$ and $\rho_F(x) = 0$ for $x \geq x_0$. This is in excellent agreement with the steady state ρ_F from our numerical model with no stochasticity and only cytoplasmic transport, as shown in Figure 3.3(a).

In our non-stochastic model all cells with free fixed nitrogen levels above zero grow at the same rate and so incorporate fixed nitrogen at the same rate. This produces
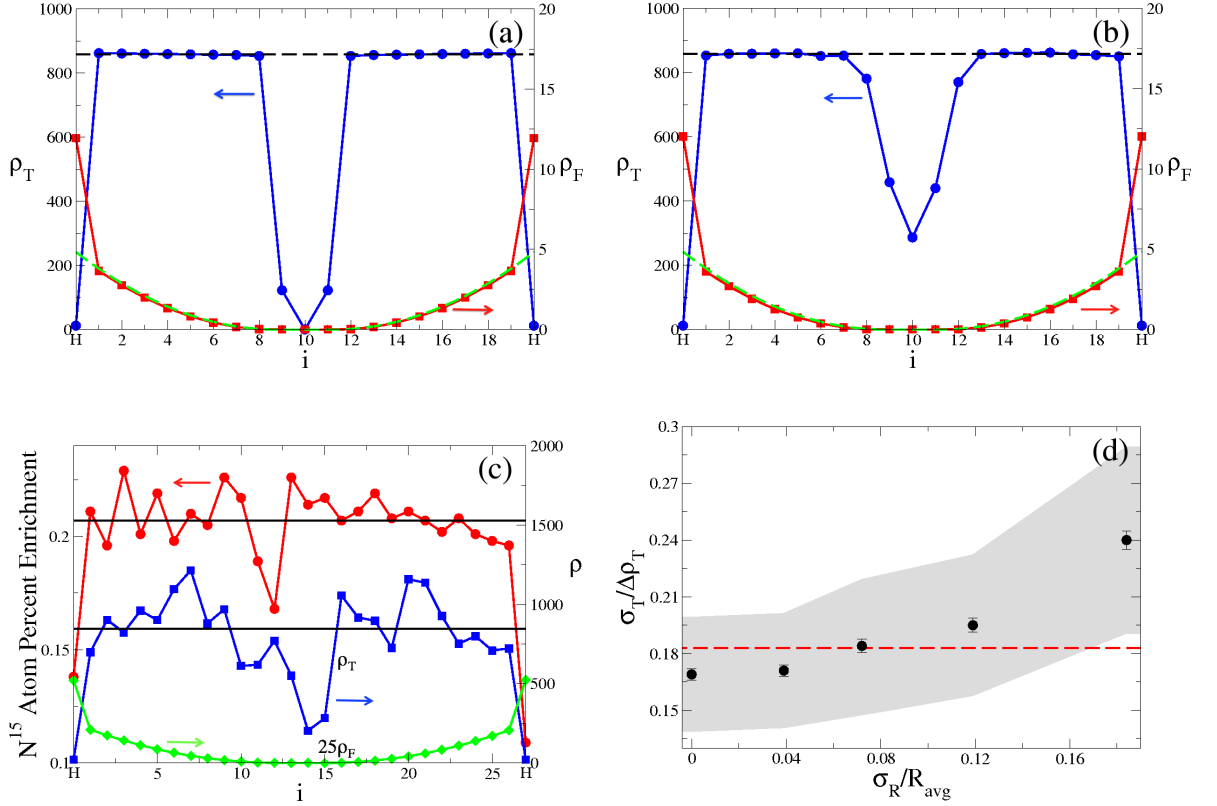
Figure 3.3: (a) non-stochastic concentrations after $t = 4$h of labeled fixed nitrogen exposure vs. the cell index i. Heterocysts are indicated by ("H"). Shown are the total fixed nitrogen (ρ_T, blue solid line with circles using the left axis) and the free fixed nitrogen (ρ_F, red solid line with squares using the right axis). Note the dramatic difference in scale. The steady state analytic result for the free fixed nitrogen from a continuum approximation, Equation 3.6, is shown as a green line; while the corresponding plateau in ρ_T due to incorporated fixed nitrogen is shown by the black dashed line. (b) averaged stochastic ρ_T using doubling-time variation $\Delta = 4.5$h, with error bars too small to be seen outside data points. The average is over 1000 distributions. (c) the experimental fixed nitrogen concentrations of Popa *et al* [8] in red circles (scale on left axis) and from a single stochastic result from our model using Δ = 4.5h and $G_{het} = 4.25 \times 10^6$ s^{-1} with ρ_T in blue squares and $25\rho_F$ in green diamonds (scale on right axis, offset for clarity). Also shown as black lines are the vegetative cell averages. (d) the coefficient of variation in vegetative cell fixed nitrogen levels, $\sigma_T/\Delta\rho_T$, vs. the coefficient of variation in the growth rate of the vegetative cells, σ_R/R_{avg}. The error bars indicate statistical errors of the average variation measured over 100 samples. The grey shaded region indicates the size of the standard deviation, which corresponds to the expected error for a single sample. The red dashed line indicates the variation in the experimental data of Popa *et al* [8], with a single sample. For (a)-(d), we have used only cytoplasmic transport with $D_C = 1.54$ μm s^{-1} and $D_I = D_E = D_P = 0$.

a constant plateau in ρ_I for vegetative cells between the heterocyst at $x = 0$ and the cell where the free fixed nitrogen vanishes, i.e. for $0 < x < x_0$. We see this in the discrete cellular model, which is shown with ρ_T in Figure 3.3(a). In this figure there are heterocysts at both ends of the distribution, indicated by "H", and vegetative cells are indexed by i, counting from the left heterocyst. The cells $i = 9$ and $i = 11$ have significantly less fixed nitrogen than cells 1-8 and 12-19. This is because the free fixed nitrogen ρ_F never goes above zero for cells 9 and 11 and they are only able to grow using Φ_{in} as described in Section 3.2.2. All the other cells reach non-zero ρ_F and grow at their optimum rate. The height of the plateau reflects how much time has elapsed since the labelled fixed nitrogen was provided.

In Figure 3.3(a) we show non-stochastic results for $t = 4$h, to be able to directly compare with the data of Popa [8]. The predicted plateau height ($<R^{opt}>gt/<L>$) from our previous continuum model is shown with a black dashed line, and the agreement is excellent. We also see that the free fixed nitrogen ρ_F (scale on right side) is approximately two orders of magnitude smaller than the total fixed nitrogen ρ_T (scale on left side), which is dominated by the incorporated fixed nitrogen ρ_I. We see excellent qualitative agreement with the plateau-like distribution seen by Popa. The gradient of free fixed nitrogen on top of the much larger amount of total fixed nitrogen is too small to be easily resolved by the experimental technique.

3.3.3 Stochasticity of long time distributions

Figure 3.3(b) shows the total fixed nitrogen distribution using our stochastic model that has been averaged over 1000 vegetative segments between two heterocysts. It looks qualitatively very similar to the deterministic results of Figure 3.3(a). A direct comparison of a single ρ_T distribution produced by our stochastic model (blue curve with squares, right axis) and the experimental data of Popa (red curve with circles, left axis) is in Figure 3.3(c) (the two curves are offset by an arbitrary amount). The distributions are qualitatively quite similar and much rougher than those of Figures 3.3(a) and (b). However, the ρ_F from the stochastic model (green curve, diamonds) is very smooth, maintaining the gradient despite the noisy ρ_T distribution and indicating that the roughness can entirely be attributed to incorporated fixed nitrogen, ρ_I.

In our stochastic simulation, the variation of ρ_T shown in (c) is due both to the

varying growth rates of cells along the filament and to their varying lengths. This leads to the observed differences in the density of incorporated fixed nitrogen after a fixed time. This means that it is possible that the noise in the distribution seen by Popa is due to the stochasticity of the individual cell's growth rates, rather than the experimental apparatus or method. In (d) we plot the coefficient of variation of the total fixed nitrogen, $\sigma_T/\Delta\rho_T$ against the coefficient of growth rate variation, σ_R/R_{avg}. Here σ_R is the standard deviation of the growth rate distribution, and R_{avg} is the average growth rate. Similarly σ_T is the standard deviation of ρ_T in vegetative cells, while $\Delta\rho_T$ is the difference between the average ρ_T in vegetative cells and the average ρ_T in the heterocysts. Because we have implemented the growth rate range in our model, we are able to calculate σ_R/R_{avg} exactly for our simulated filaments. For $\sigma_T/\Delta\rho_T$ for our simulated filaments in Figure 3.3(d) we measured over 100 filaments to get an average (shown as black circles), standard deviation (shown as the grey region), and standard deviation of the mean (black error bars on black circles). The red dashed line represents the same quantity for the experimental Popa data when determined by the same procedure as for our simulation data. Given that the experimental data is for a single sample, the experimental value is best compared to the single standard deviation region in grey. It shows that the fixed nitrogen distribution variability for a broad range of growth rate variabilities is consistent with the observed variability in the Popa nitrogen distribution. This range is near the coefficient of variation $\sigma_R/R_{avg} = 0.165$ reported for a mutant of PCC 7120 [56]. We conclude that the variability in the Popa nitrogen distribution may be entirely due to stochasticity of the individual cell growth rates.

After extended isotopic dinitrogen exposure, much longer than typical cell doubling times, we expect a nearly constant distribution of ρ_T — corresponding to the nearly uniform concentration of fixed nitrogen in a cyanobacterial filament. Notably, the variation within the plateau, evident in the data of Popa *et al* [8] and modelled in Figure 3.3(d), will also vanish at late times. After a long enough period of time, most of the fixed nitrogen in the filament will be labelled and among vegetative cells it will be evenly distributed. Old heterocysts will retain their dip for longer than vegetative cells retain their variability, but nitrogen recycling will occur through basal processes and eventually even the old heterocysts will have the same labelled nitrogen density

as the rest of the filament. On top of the large and nearly constant ρ_I in the cells will be the small gradient of ρ_F that is described in Equation 3.6 and illustrated in Figures 3.3(a)-(c). It is possible that determining the fixed nitrogen distribution in filaments that have been grown for weeks would allow the small gradients of ρ_F to be discerned with NanoSIMS techniques [8].

3.4   Periplasmic transport

3.4.1   Long time nitrogen distributions with cytoplasmic and periplasmic transport

We have established that the qualitative characteristics of both the spiked distribution of Wolk *et al* [11] and the flat distribution of Popa *et al* [8] can be reproduced using a single quantitative model for the cyanobacterial filament using only cytoplasmic transport. In this section, we explore how the addition of periplasmic transport affects the Popa-like plateau formed in ρ_T.

In Figures 3.4(a) and (b) import and export between the cytoplasm and periplasm has $D_I = 4.98$ s^{-1} and $D_E = 0.498$ s^{-1} (see Section 3.2.1). The only parameter that varies is the value of the periplasmic diffusivity D_P, as indicated by the legend in (b). We vary D_P up to the magnitude expected for unimpeded diffusive transport of glutamine along the filament, as discussed in Section 3.2.1. The characteristic dip in the centre of the total fixed nitrogen distribution, ρ_T, in (a) fills in at larger values of D_P, but is not sufficient to qualitatively change the form of the distribution. The mid-segment dip fills in because periplasmic transport allows free fixed nitrogen to move past the first starving vegetative cell with $N_C = 0$. To evaluate how significant the periplasmic transport is, we show in (b) the fraction of the flux along the periplasm to the total flux along the cytoplasm and periplasm. This essentially shows how much of the fixed nitrogen is transported through the periplasm. The flux is mostly cytoplasmic next to the heterocysts, and mostly periplasmic next to starving cells. In between, up to 50% of the flux can be periplasmic. For the larger D_I and D_E values of Figures 3.4(c) and (d) the results are very similar. These results indicate that, within our model, the periplasmic transport cannot be the dominant form of transport of fixed nitrogen from heterocysts to vegetative cells. However, we have shown that a
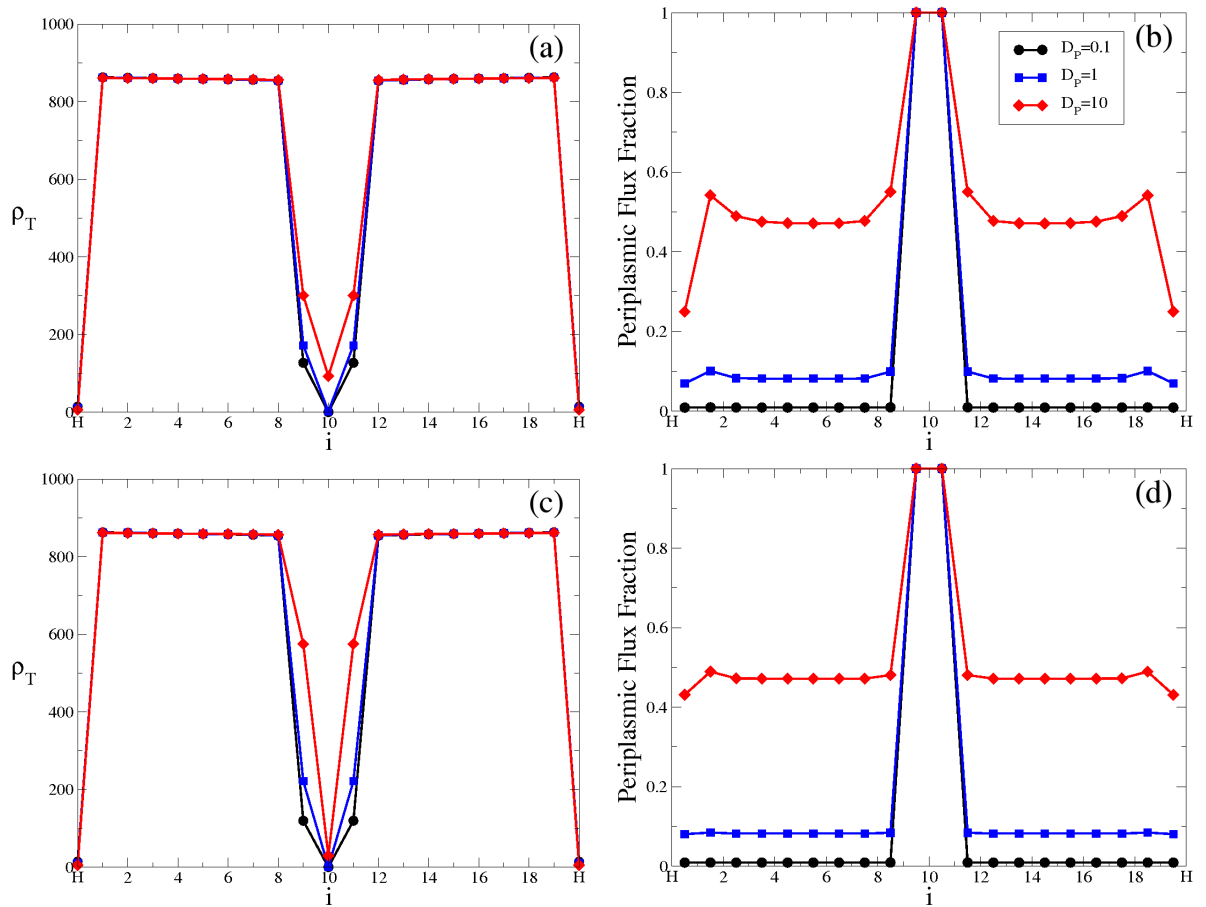
Figure 3.4: Fixed nitrogen distributions for systems with additional periplasmic transport, as discussed in Section 3.4.1. In all four plots the periplasmic transport rate D_P (in units of μm s^{-1}) is varied according to the legend in (b). (a) the total fixed nitrogen distribution, ρ_T, for transport between the cytoplasm and periplasm with $D_I = 4.98$ s^{-1} and $D_E = 0.498$ s^{-1}. The data of (b) corresponds to the curves of the same colour in (a) and shows the fraction of the total fixed nitrogen flux along the filament that is periplasmic vs. the cell index i. (c) and (d) are similar to (a) and (b), respectively, except that import and export between the cytoplasm and periplasm are tenfold stronger, with $D_I = 49.8$ s^{-1} and $D_E = 4.98$ s^{-1}. As before, D_C=1.54 μm s^{-1}. We use $D_L = 0.0498$s^{-1} and $G_{het} = 3.67 \times 10^6$ s^{-1}.

range from completely cytoplasmic transport, to roughly half cytoplasmic and half periplasmic transport is consistent with the qualitative fixed nitrogen distribution reported by Popa *et al* [8]. Nevertheless, cytoplasmic transport alone can be used to determine the fixed nitrogen distribution and the growth it controls.

3.4.2 Long time nitrogen distributions with only periplasmic transport

We have also considered the model in absence of cytoplasmic transport, with only periplasmic transport of fixed nitrogen.

In Figure 3.5(a), import and export between the cytoplasm and periplasm has D_I = 4.98 s^{-1} and D_E = 0.498 s^{-1}. Again, we have only varied D_P, as indicated by the legend in (b). With only periplasmic transport in the filament, there is a very different fixed nitrogen distribution. For the slowest transport, $D_P = 0.1$ μm s^{-1}, only the two cells nearest the heterocysts are growing at their maximum rate, with the third receiving a much smaller amount of nitrogen. Cells further than three cells from the heterocyst are receiving no nitrogen and are not growing. As the periplasmic transport rate increases, fixed nitrogen reaches cells further from the heterocysts. For $D_P = 10$ μm s^{-1} some fixed nitrogen has reached all of the cells, although we still do not observe a Popa-like dip. Instead there is a broad trough in the middle of the distribution, indicating that fixed nitrogen in the periplasm is able to get past the first starving cell (with $N_C = 0$) and somewhat supply cells even further from the heterocyst. Also note that the height of ρ_T does not change with how many cells are being supplied with nitrogen. For slower transport, the nitrogen that is not being used for growth is being lost through leakage to outside the filament before it can be transported to starving cells.

In Figure 3.5(b), import and export between the cytoplasm and periplasm are tenfold faster, at $D_I = 49.8$ s^{-1} and $D_E = 4.98$ s^{-1}. The story here is similar to (a), where for slower transport the middle cells are starved, and as the transport is sped up, more cells are reached. In contrast, a broad trough is never achieved and the boundary between non-starving cells and starving cells remains sharp. This is the effect of increasing the import and export by a factor of ten: fixed nitrogen is unable to sneak past the first starving cell it encounters. It also suggests that for some value of the periplasmic transport rate coefficient, D_P, the distribution would exhibit
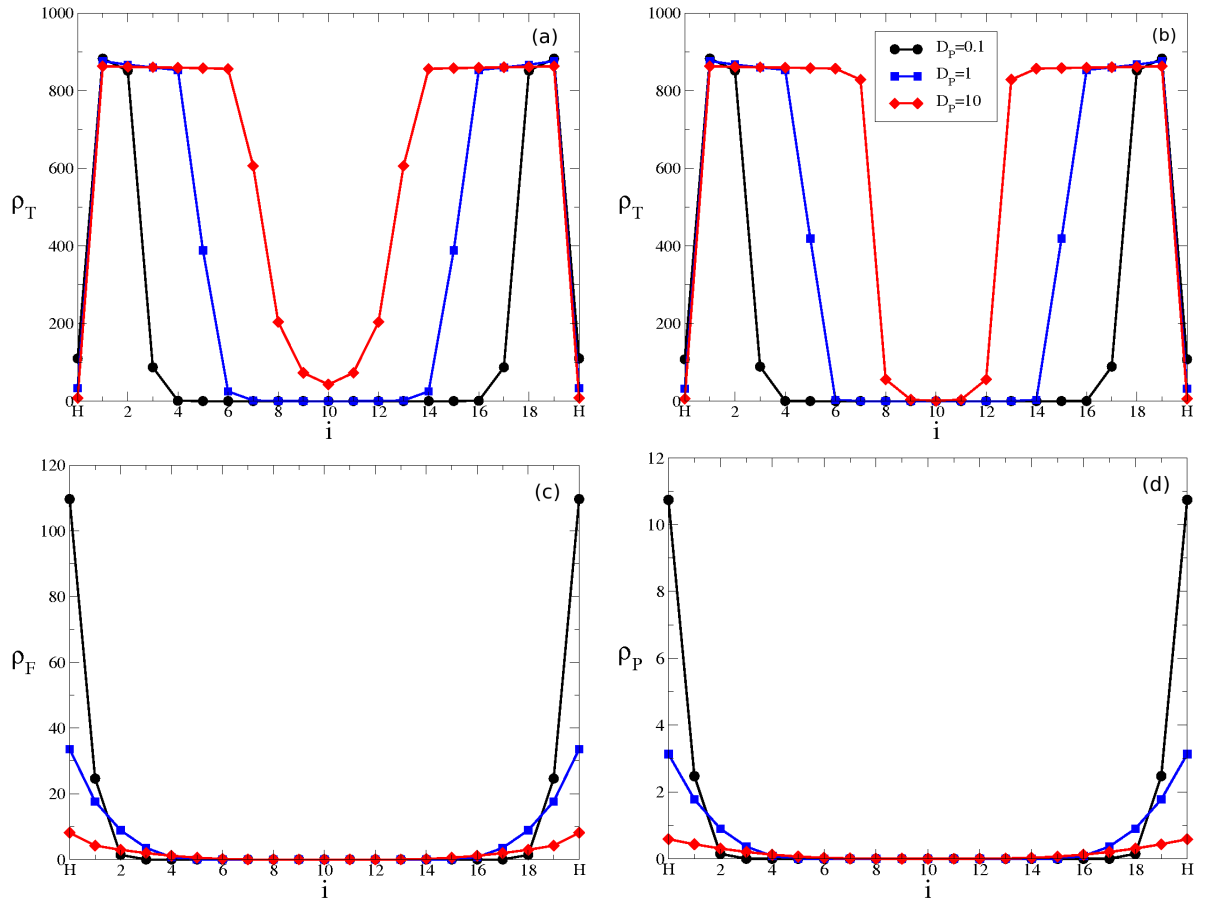
Figure 3.5: Fixed nitrogen distributions for systems with only periplasmic transport $(D_C = 0)$, as discussed in Section 3.4.1. In all four plots the periplasmic transport rate D_P (in units of μm s^{-1}) is varied according to the legend in (b). (a) the total fixed nitrogen distribution, ρ_T, for transport between the cytoplasm and periplasm with $D_I = 4.98$ s^{-1} and $D_E = 0.498$ s^{-1}. (b) the total fixed nitrogen distribution, ρ_T, for transport between the cytoplasm and periplasm with $D_I = 49.8$ s^{-1} and $D_E = 4.98$ s^{-1}. (c) the cytoplasmic fixed nitrogen distribution with $D_I = 4.98$ s^{-1} and $D_E = 0.498$ s^{-1}. (d) the periplasmic fixed nitrogen distribution with $D_I = 4.98$ s^{-1} and $D_E = 0.498$ s^{-1}.

a Popa-like dip. We have shown that cytoplasmic transport, at the experimentally measured rates, is able to do this. A similar measurement of the rates associated with periplasmic transport would be necessary to assert this for periplasmic transport as well.

In Figures 3.5(c) and (d) we show the free cytoplasmic (ρ_F) and the periplasmic (ρ_P) fixed nitrogen distributions respectively, for the case of $D_I = 4.98$ s^{-1} and $D_E = 0.498$ s^{-1}, which is shown in (a). Cells furthest from the heterocyts do not have free fixed nitrogen in the cytoplasm for any rate of periplasmic transport, but the fastest transport rate allows some fixed nitrogen to be in the periplasm for cells furthest from the heterocysts. The loss from the filament is proportional to N_P, and because ρ_P is so high for slow transport, more fixed nitrogen is lost via leakage and it reaches fewer cells.

3.4.3   Long time nitrogen distributions with graded growth

In Section 3.1 we mentioned the growth model of Wolk *et al* [11], which assumed that growth was linearly related to fixed nitrogen concentrations. Our growth model, laid out in Section 3.2.2, instead has the growth of a cell be constant unless there is not enough free fixed nitrogen to support it. All the results in the above sections in this chapter have used our constant growth model. Here we will address the linear or graded growth model. Growth is expected to increase when nutrient concentration increases, in contrast to our model. We want to show that a graded model cannot explain the nitrogen distribution of Popa *et al* [8].

With the graded growth that is linearly proportional to ρ_F, the dynamics or ρ_F are then

$$\partial \rho_F / \partial t = D\nabla^2 \rho_F - k\rho_F, \tag{3.7}$$

The first term on the right represents diffusion and the second is a local sink due to vegetative growth. Comparing to Equation 3.5 we see that this second term has changed from a constant term in Equation 3.5 to a linear term here. As above with Equation 3.5 we ignore the discreteness of cells so that we can obtain an analytic solution. Following Section 3.3.2, a heterocyst at $x = 0$ has a flux of $J = -D\partial \rho_F / \partial x = G_{het}/2$ to both sides. Demanding a steady state, $\partial \rho_F / \partial t = 0$,

and applying the heterocyst flux boundary condition yields the solution

$$\rho_F = \frac{1}{2\sqrt{kD}}e^{-\sqrt{k/D}x}. \tag{3.8}$$

The amount of free fixed nitrogen used to grow by the filament and therefore incorporated by the cell is proportional to the free fixed nitrogen concentration, ρ_F, and therefore the total fixed nitrogen, ρ_T, will be proportional to ρ_F and decay away exponentially from the heterocyst with a length scale $\sqrt{D/k}$. This is inconsistent with the nitrogen distribution of Popa et al [8] for all $\sqrt{D/k}$. A large $\sqrt{D/k}$ will lead to a gradual decrease in the ρ_T at larger x while a small $\sqrt{D/k}$ will lead to a rapid decrease in ρ_T at small x.
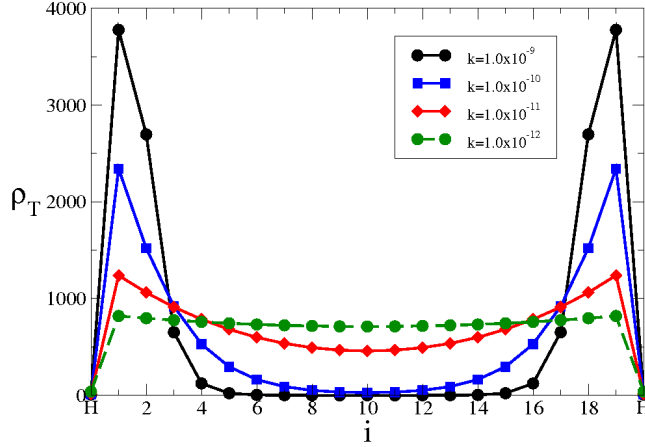
Figure 3.6: The numerical total fixed nitrogen distributions vs. cell index i using graded growth proportional to local ρ_F as Eq. 3.9. The growth rate proportionality constant k (in units of μm s^{-1}) is varied, as indicated by the legend. This is a deterministic cellular model with cytoplasmic transport: $D_C = 1.54$ μm s^{-1} and $D_I = D_E = D_P = 0$.

With our discrete deterministic model, we have modelled the graded growth

$$R_{graded}(N_c) = kN_c/L_i. \tag{3.9}$$

This equation can be used just as Equation 3.4 was used in Equation 3.3 to determine the value of the sink term G in Equation 3.1. The results are shown in Figure 3.6. We either recover a plateau with no dip in the middle region (for small k) or a quick drop to a central trough with no plateau. The discrete nature of the model modifies the observed pattern near bounding heterocysts compared to Equation 3.8.

3.5   Discussion

We have presented a quantitative model for fixed nitrogen transport and dynamics in a cyanobacterial filament, including growth of vegetative cells and production of fixed nitrogen by heterocysts. The goal of this model in this chapter was to understand how the fixed nitrogen transport in the filament determines the fixed nitrogen distribution, and to use this understanding to address the disagreement of the nitrogen distributions of Wolk *et al* [11] and Popa *et al* [8]. We also wanted to investigate how the distribution changed when the model included cytoplasmic and/or periplasmic connections that could transport fixed nitrogen.

Our model has reproduced the qualitative fixed nitrogen distribution near heterocysts seen after short exposure to fixed nitrogen (Figure 3.2), similar to the spiked distribution of Wolk *et al* [11], as well as the distribution seen after much longer exposure to fixed nitrogen (Figure 3.3), similar to the plateau distribution of Popa *et al* [8]. These two distributions, seemingly at odds, are really the product of the same dynamics but at different time scales. The apparent lack of a fixed nitrogen gradient away from the heterocysts in the Popa distribution is explained by the much larger amount of incorporated fixed nitrogen, which also explains the noisy plateau between heterocysts. When modelled, the free fixed nitrogen does show a smooth gradient away from the heterocysts even when the total fixed nitrogen is quite rough (Figure 3.3(c)).

Qualitatively, our results indicate the importance of the large, noisy, plateau-shaped concentration profile of incorporated fixed nitrogen. The level of the plateau will increase linearly in time after the free fN gradient away from heterocysts reaches a steady state, which takes minutes. The free fN, ρ_F, is soon much smaller than the incorporated fN, ρ_I. The spike in the Wolk distribution, seen in the first few minutes of exposure to isotopically labeled dinitrogen, was when the free fN still dominated the total ρ_T. In contrast, the plateau seen in the Popa distribution was after hours of exposure, when the total fN was dominated by the incorporated component and the gradient in ρ_F was relatively small even with respect to variations of ρ_I within the plateau (see Figure 3.3(c)). At times that are long after the steady state is reached but still much shorter than division times the heterocysts present dips of labelled fixed nitrogen, as observed by Popa *et al* [8]. Stochasticity, through initial cell lengths and

cellular growth rates, adds noise to the incorporated fN distribution, ρ_I. As we show in Figure 3.3 (d), the modelled cell-to-cell variation is due to differences in growth rate which manifest themselves as different amounts of labeled fN incorporated by growth. We believe this explains the observed variation in the experimental nitrogen distribution of Popa $et\ al$ [8].

The possibility of periplasmic fixed nitrogen transport raised by a continuous periplasm [82], transport of GFP [83], and an outer membrane permeability barrier [86] led us to explore the possible impact of periplasmic transport in addition to cytoplasmic cell-to-cell connections [7]. As illustrated in Figure 3.4, we found that a significant proportion (up to 50%) of fN transport along the filament can occur through the periplasm without qualitatively changing our fN distributions. We cannot rule out a role for the periplasm in fN transport, and larger proportions may be achievable with more extensive parameter searches. Indeed, the fact that knocking out amino acid permeases leads to impaired diazotrophic growth [60, 87, 88] does indicate that significant leakage from the cytoplasm does occur (i.e. $D_E > 0$) — which is a necessary part of periplasmic transport. Nevertheless, the fact that modelling with only cytoplasmic transport was able to qualitatively match both experimental distributions means that our results are also consistent with no significant periplasmic transport. For simplicity, the models developed in future chapters will be without periplasmic transport.

We have also explored the resulting distribution when there is only periplasmic transport and there is no cytoplasmic transport ($D_C = 0$). The results are in Figure 3.5. For relatively slow import and export between cytoplasm and periplasm, they show either a large trough lacking fixed nitrogen between the heterocysts, or a wider dip that does not reach zero. For relatively fast import and export between the cytoplasm and periplasm, they show a sharp transition from a constant fixed nitrogen level to zero. It seems possible that with fine tuning to some intermediate rate of import and export, the fixed nitrogen distribution could match the Popa plateau phenotype. With these results, it seems very unlikely that fixed nitrogen is transported through the periplasm alone.

We have used a simple model, Equation 3.4, to limit growth in response to local fixed nitrogen starvation: cells grow at a fixed rate R^{opt} if there is sufficient free fixed

nitrogen available locally ($N_C > 0$), and at the maximal rate allowed by local input fluxes of free fixed nitrogen if not. This model reflects the absolute limitation on growth placed by available fixed nitrogen, but also recovers the characteristic plateau of ρ_T seen in the Popa distribution. As we illustrate in Section 3.4.3, "graded" growth rates, that smoothly depend on N_C and vanish at $N_C = 0$, respond to the graded distribution of free fixed nitrogen away from heterocysts, Equation 3.8, and produce graded growth patterns and graded distributions of ρ_I and ρ_T (see Figure 3.6), unlike those observed by Popa [8]. Graded growth models, such as the Monod model [100], are typically designed to describe response to external metabolite concentrations rather than cytoplasmic concentrations. Like our growth model, *Salmonella typhimurium* [101] and *Escherichia coli* [102, 103] both show non-graded growth in response to cytoplasmic fixed nitrogen starvation. However, there is no direct evidence of how growth of vegetative cells in cyanobacterial filaments depends on the cytoplasmic freely available fixed nitrogen, ρ_F.

The experimental work was done on different species of filamentous cyanobacteria, namely *Anabaena cylindrica* by Wolk *et al* [11] and *Anabaena oscillarioides* by Popa *et al* [8]. In contrast, our models were parameterized by consideration of transport studies done in *Anabaena* sp. strain PCC 7120 (see Section 3.2.1). While we expect parameter values to differ between species, so that precise numerical agreement with the Wolk and Popa results is not to be expected, our qualitative agreement indicates that similar fixed nitrogen transport and dynamics could apply in these different model organisms.

We expect the qualitative features of the Wolk and Popa fixed nitrogen distributions to apply to the early and late time fixed nitrogen distributions, respectively, of all filamentous cyanobacteria with heterocysts that are growing in media lacking fixed nitrogen and that have significant cytoplasmic connections. Nevertheless, quantitative details will depend, via the model parameterization, on the cyanobacterial strain and the experimental conditions. For example, sufficiently large periplasmic transport (via D_P) would broaden Popa's dip in incorporated fixed nitrogen between heterocysts due to transport and growth past the first starving cell. Conversely, a larger cytoplasmic diffusivity (via D_C) would alter the magnitude of Wolk's free fixed nitrogen (via Equation 3.6) but not the incorporated fixed nitrogen. The average

doubling time determines the characteristic time at which we expect the later time distribution to emerge. Cell to cell variations in the doubling time, in turn, determine the amount of variation expected within the Popa-like plateau of incorporated fixed nitrogen near heterocysts for the first few generations of labelled growth.

Our results reconcile the metabolite patterns shown by the autoradiographic technique of Wolk *et al* [11] with the NanoSIMS technique of Popa *et al* [8]. It is not obvious which would currently have the best resolution, and it would be useful to have either or both applied to *Anabaena* PCC 7120 — the most popular current model for filamentous cyanobacteria and for which recent transport studies exist [7, 83]. We can make four qualitative but testable predictions from our work. First, as the duration of labelled fN is increased the total fN distribution away from heterocysts will evolve from a Wolk-like peak, dominated by free fN, to a Popa-like plateau, dominated by incorporated fN. As indicated by Figure 3.2, the crossover takes place after only a few minutes of labelled dinitrogen fixation. Second, on a similar timescale the free fN will reach a steady state smooth parabolic shape even in a single sample. This free fN may be directly measurable in vivo, for example, by fluorescence resonance energy transfer (FRET)-based metabolite nanosensors [104]. The parabolic gradient of ρ_F could also be uncovered by averaging the results of many experiments measuring ρ_T. Third, the variations seen by Popa *et al* [8] will decrease in relative magnitude after at least one doubling time of exposure to labeled fN. Fourth is more of an assumption, but we expect that growth of vegetative cells will slow abruptly when the freely available fN vanishes — according to Equation 3.4 rather than in a graded fashion according to Equation 3.9. Finally, with more precise parameterization of our transport parameters, sources, and sinks we can make quantitatively testable predictions. Prediction of the variance of cell-to-cell fN, σ_I, requires only the observed growth rate variance σ_R and the duration of exposure to labeled dinitrogen.

Chapter 4

Growth Rates of and Heterocyst Placement in the Cyanobacterial Filament

4.1 Motivation and background

This chapter is based on a paper we have published in Physical Biology, volume 9, page 046002, 2012 [105]. IOP Publishing has given me permission to reproduce the contents of the paper in this thesis (see Appendix B). For this paper my contributions were to write the simulation, do the analytical and computational work, generate the figures, and write the first draft. I was also an equal partner with my supervisor in developing the model and revising the paper.

Chapter 3 has allowed us to understand the dynamics of fixed nitrogen in the cyanobacterial filament. It has established, by comparison with experimental nitrogen distributions, that our model is a reasonable approximation of the actual fixed nitrogen dynamics of a living filament. It has demonstrated that a model including cytoplasmic transport alone is sufficient to be consistent with the experimental nitrogen distributions. It has also shown that our constant growth model, in contrast to a graded growth model, can explain the distribution of nitrogen between two heterocysts.

With our fixed nitrogen model established, we wanted to investigate the growth of filaments on time scales larger than division times, which also facilitates investigation of the dynamic placement of heterocysts. In comparison to Chapter 3 we have made several changes to the model. We now only consider the cytoplasmic compartment, and no longer consider the periplasmic compartment. We add external fixed nitrogen and its import. We now need to dynamically place heterocysts instead of only having two heterocysts at the beginning of a simulation. Detailed explanations of these changes are in Section 4.2.

As detailed in Section 2.4, significant progress has been made in characterizing the

genetic network underlying the observed heterocyst pattern [3, 15] that is developed in response to the absence of external fixed nitrogen. The efforts to understand the genetic network and other possible mechanisms controlling heterocyst differentiation focus on understanding how the cells in the filament choose which particular cells will be heterocysts. While we will address heterocyst selection from a physical point of view, we will focus on understanding why the heterocyst pattern exists and what advantages it provides. Importantly, we wanted to examine growth not just at very high or nearly zero exogenous fixed nitrogen, but at intermediate levels.

As we shall show, a filament with too few (or inactive) heterocysts will starve of fixed nitrogen and grow slowly without external fixed nitrogen [106], while too many heterocysts, which do not grow and divide, will also inhibit growth [47]. Balancing these effects in the filament qualitatively explains the heterocyst frequency of approximately 10% observed with no external fixed nitrogen, but does not explain the distinctive pattern of heterocyst spacings that is seen in the model cyanobacterium *Anabaena* sp. PCC 7120 [6, 26, 46, 47]. Indeed, the observation that mutant strains exhibiting a distinct pattern of multiple contiguous heterocysts (Mch) show reduced growth [62] indicates that heterocyst placement is important. In this chapter, we use quantitative modelling to explore the hypotheses that simple heterocyst placement strategies can affect filament growth, and that observed heterocyst patterns reflect placement strategies that maximize growth.

Previous modeling efforts relating to this work are reviewed in Section 2.7.

We hypothesize that leakage of fixed nitrogen from the cyanobacterial filament may distinguish between heterocyst patterns, in terms of growth rates. Evidence for leakage was first reported by Fogg *et al* [2, 107], who found fixed nitrogen products outside the filament. Supporting this, Paerl [93] observed bacteria clustering around cyanobacterial filaments, particularly at junctions between heterocysts and vegetative cells, and indicated a possible symbiosis based on enhanced nitrogen fixation and leaked fixation products. Thiel [94] found protein proteolysis byproducts in the extracellular medium after fixed nitrogen starvation. Significant leakage is also consistent with the observation of reduced diazotrophic growth in strains with impaired amino acid uptake transporters [60, 88]. Nevertheless, we are not aware of any quantitative measurements of the rate of leakage. Qualitatively, a regular pattern of heterocysts

would minimize the distance traveled by fixed nitrogen products so as to minimize leakage from vegetative cells, leaving more fixed nitrogen available for growth.

In this chapter, we examine heterocyst frequencies that maximize filament growth within the context of a quantitative transport model [75] that incorporates fixed nitrogen transport, vegetative cell growth, and fixed nitrogen production at heterocysts. We explore the impact on growth of different heterocyst placement strategies, including random placement, and find that they are almost indistinguishable in the absence of leakage of fixed nitrogen from the filament, but clearly distinct with leakage. We find that the heterocyst spacing patterns corresponding to maximal filament growth are qualitatively similar to those seen experimentally.

The genera *Anabaena*, which includes heterocystous cyanobacteria, is widely distributed geographically in freshwater lakes [108, 109], where fixed nitrogen is one of the major substances limiting growth [110, 111, 112]. Nitrification turns ammonia into nitrite and then nitrate [109, 113]. Nitrate levels in lakes range from 0-10 mg/l ($0\text{-}10\times10^{22}$ m^{-3}) in unpolluted freshwater, but vary both seasonally and spatially [109, 114]. Heterocystous cyanobacteria are also found in the oceans [115, 116] where nitrate levels have been recorded to vary from approximately 0-50 μmol/kg [117] ($0\text{-}3\times10^{22}$ m^{-3}). It has long been known that sources of fixed nitrogen initially present in the medium can increase the mean spacing between heterocysts [2, 118] and that heterocysts will differentiate at non-zero levels of external fixed nitrogen (efN) [2]. Field studies of heterocystous cyanobacteria also show both a range of heterocyst counts and a range of efN concentrations, with significant positive correlations between the two (see e.g. [119]). In addition, steady state chemostat experiments show that *Anabaena flos-aquae* can adjust nitrogen fixation to achieve constant growth, within approximately 10%, as efN concentrations are varied [29]. While growth independence from efN seems desirable in the face of environmental variability, it raises the question of how and how well it is achieved in terms of heterocyst fraction and pattern.

The heterocyst pattern changes with time as it evolves towards a steady state distribution after efN deprivation [26, 46, 47]. Earlier models [31, 74] have focused on the early heterocyst pattern observed 24 h after efN deprivation. While the early pattern and the later patterns are qualitatively similar, with broad distributions of

heterocyst spacings ranging from zero to more than twenty cells between heterocysts with a peak at about a ten cell spacing, we focus on the steady state distribution in this paper. We explore the hypothesis that cyanobacterial filaments use a unified heterocyst placement strategy even at non-zero levels of efN. With all of our strategies, we find that maximal growth is observed with heterocyst frequencies that decrease continuously with increasing levels of efN.

4.2 Model

4.2.1 Fixed nitrogen transport

Our model of fixed nitrogen (fN) transport and incorporation is similar to the model from Chapter 3. Since periplasmic transport was not required to explain the experimentally measured nitrogen distributions, we will not be including periplasmic nitrogen in our model here. Accordingly, any leakage of fixed nitrogen from the filament will be modelled as being directly lost from the cytoplasm. We are considering exogenous sources of fixed nitrogen and therefore also have import of fixed nitrogen from the extracellular medium. This chapter also does not focus on accumulation of nitrogen, so we track only the freely diffusing fixed nitrogen of Chapter 3 and call it just nitrogen, with the understanding that this is the cytoplasmic freely diffusing fixed nitrogen. Our model tracks the amount of fN, $N(i,t)$, in each cell i vs. time t:

$$\frac{d}{dt}N(i,t) = \Phi(i,t) + D_O \rho_{efN} l(i,t) - D_M N(i,t) + G_i, \qquad (4.1)$$

Φ is the same as Φ_C in Equation 3.1 that is described by Equation 3.2.1. We will still use $D_C = 1.54$ μm s^{-1} between two vegetative cells and $D_C = 0.19$ μm s^{-1} between a vegetative cell and a heterocyst and do not change these values in this chapter. D_O is the coefficient of the term describing import of fixed nitrogen from outside the filament, and D_M is the coefficient of the term describing loss of fixed nitrogen to the medium outside the filament. These transport processes are shown schematically in Figure 4.1.

To estimate D_O, we use Fogg's observations of heterocyst frequency in a medium to which a fixed amount of efN in the form of ammonia was added [2]. Heterocyst frequency decreased in time due to vegetative growth until the ammonia concentration dropped to 4×10^{-5}M (see p. 245 of [2]). We assume that a slightly higher
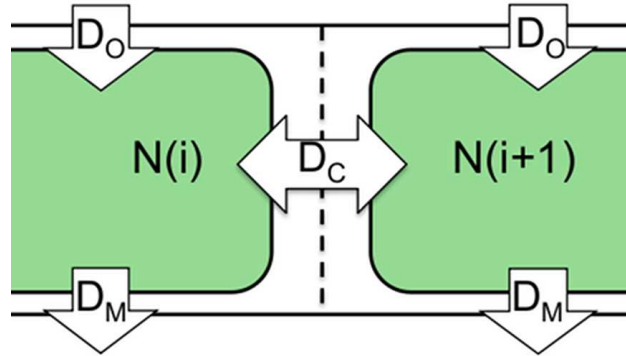
Figure 4.1: Schematic of fixed nitrogen transport as represented by Equation 4.1. $N(i)$ is the amount of freely diffusing fixed nitrogen for cell i. D_C governs transport between cells, D_O governs import from the medium outside the filament into the cell, and D_M controls loss from the cell to the external medium.

concentration of efN, 5×10^{-5}M ($\rho_{efN} = 3 \times 10^{22}$m^{-3}), is just sufficient for maximal growth. From Equation 4.1 the amount of fixed nitrogen imported from the external medium in a time T is $D_O \rho_{efN} l T$. For a cell to double in length over a period of time $T = 20$ h it needs approximately 1.4×10^{10} N atoms to be imported [75]. Using $l = l_{min} = 2.25$ μm, this yields $D_O = 2.9 \times 10^{-18}$m^3/($\mu$m·s).

We also include a loss term with coefficient D_M in Equation 4.1, following [2, 60, 88, 93, 94, 107]. Loss is expected to be proportional to the cytoplasmic density, $N/(Al)$, where l is the length of the cell and A is the cross-sectional area, and also to the number of transporters, which will themselves be proportional to the cell length. This results in a loss term that is independent of length l, with a coefficient D_M with units of s^{-1}. There are no direct measurements of leakage rates of fN from cyanobacterial filaments that we are aware of. We take the loss rate as small compared to the import rate, and so use either $D_M = 0.01 D_O/A$ (referred to as 1% loss) and $D_M = 0.1 D_O/A$ (referred to as 10% loss). We use a fixed $A = \pi \mu$m^2 corresponding to a radius of 1μm.

As in the model in Chapter 3, in addition to transport terms, there is also a source/sink term G_i in Equation 4.1 that describes fN production and consumption in the heterocysts and vegetative cells, respectively. This G term is discussed in the next section.

4.2.2 Cell growth and division

Our model of cell growth and division is very similar to the model in Chapter 3 described in Section 3.2.2, with some changes. Instead of selecting randomly and uniformly between a minimum and a maximum growth rate, as in Chapter 3, we select randomly and uniformly from between a minimum cell period $T_{min} = T_D - \Delta$ and a maximum cell period $T_{max} + \Delta$. We anticipate this change will not significantly affect the results. As in Chapter 3, $T_D = 20$ h and $\Delta = 4.5$ h. The cell growth rate is then $R^{opt} = l_{min}/T$. Also, the flux into the cell Φ_{in} in Equation 3.4 that determines the amount of growth has changed to $Rg = \Phi_{in} = \Phi_{L,in} + \Phi_{R,in} + D_O \rho_{efN} l_i$.

4.2.3 Heterocyst placement

While we know that experimentally observed heterocyst patterns are not due to random placement [31], it is useful to evaluate the steady state growth rate achievable with random heterocyst placement. We will use random placement as a point of reference for other heterocyst placement strategies. We consider three simple strategies: random placement, random placement with no contiguous heterocysts, and local placement. For all of them, once a heterocyst is placed it immediately stops growing ($R_i^{opt} = 0$) and fixes nitrogen ($G_i = G_{het}$). Commitment to differentiation does not occur until after approximately 8 h of efN deprivation and can take as long as 14 h [26, 27, 28]. Heterocysts mature and begin to fix nitrogen approximately 18-24 h after efN deprivation [84, 26, 27] (though see [56]). For random strategies, heterocyst placement corresponds to when heterocysts begin to produce fN. For local placement, we have an explicit delay (see below).

Our first, reference, strategy is random placement ("random"). The heterocyst fraction f is fixed, and during filament growth a random vegetative cell is replaced with a heterocyst whenever possible but without exceeding f.

Our second strategy ("no-Mch") reflects the observation that contiguous heterocysts are not observed during normal development [15, 84]. It consists of our random strategy, but with the additional restriction that vegetative cells adjacent to existing heterocysts are never selected for development.

Our third strategy is local heterocyst placement ("local"). Any vegetative cell that has been continuously starving for a defined interval τ is changed into a heterocyst.

Starvation is defined as $N(i) = 0$, so that these cells have reduced growth $R < R^{opt}$ for a significant period of time. We vary τ within the range $1 - 20$ h. Starvation occurs due to distance from heterocysts [75], but can also reflect local clusters of fast growing cells. The waiting period τ combines the periods until commitment and maturity/nitrogen fixation discussed in Section 2.4. In Chapter 5 we will introduce a local heterocyst placement model that reflects more of the knowledge of the heterocyst commitment and differentiation process. The more complicated model of Chapter 5 is not necessary in this chapter, as we only investigate heterocysts in a steady state pattern, while the model of Chapter 5 is necessary to examine the *de novo* heterocyst pattern within the first 24 hours.

4.2.4 Some numerical details

Periodic boundary conditions were used to minimize end effects. Filaments were initiated with no heterocysts and the different strategies were followed until a steady state was reached. For the local strategy, a "no-Mch" rule was followed for the first 24 hours of differentiation to reduce initial transients.

The growth rate constant μ was calculated every six simulated hours using the total length of the filament $L(t)$, where $L(t) = L(t - 6h)e^{\mu \cdot 6h}$. Heterocyst frequency was sampled every six simulated hours by dividing the number of heterocysts by the total number of cells in the filament. Heterocyst spacings were also recorded every six simulated hours. Measurements were averaged daily, and the results of ten independently seeded runs were used to determine an overall daily average and standard deviation. All data shown is for the fifth day, which exhibits steady state for the parameters explored (in comparison with data from the fourth day).

We begin each simulation with 100 cells. Most of the phenomena we investigate occurs in filaments with a growth rate constant greater than μ =0.02 h^{-1}, which would allow the filament to grow to more than 1100 cells after five days. We use a computational timestep Δt =0.01s with a simpler Euler discretization of Equation 4.1; smaller timesteps yield indistinguishable results.

For random heterocyst placement strategies, optimal heterocyst frequencies for growth without leakage in Figures 4.5(a), and the corresponding growth rates in Figures 4.5(d), were found by locally using the standard Marquardt-Levenberg fit

algorithm with a quadratic function to the left of the optimum frequency and a linear function to the right, with the two functions meeting at the optimum frequency. Optimum heterocyst frequencies with leakage in Figures 4.7(b) and (d), and the corresponding growth rates in Figures 4.7(a) and (c), were found by a least squares quadratic fit near the maximum growth rate.

4.3   Results with no heterocysts
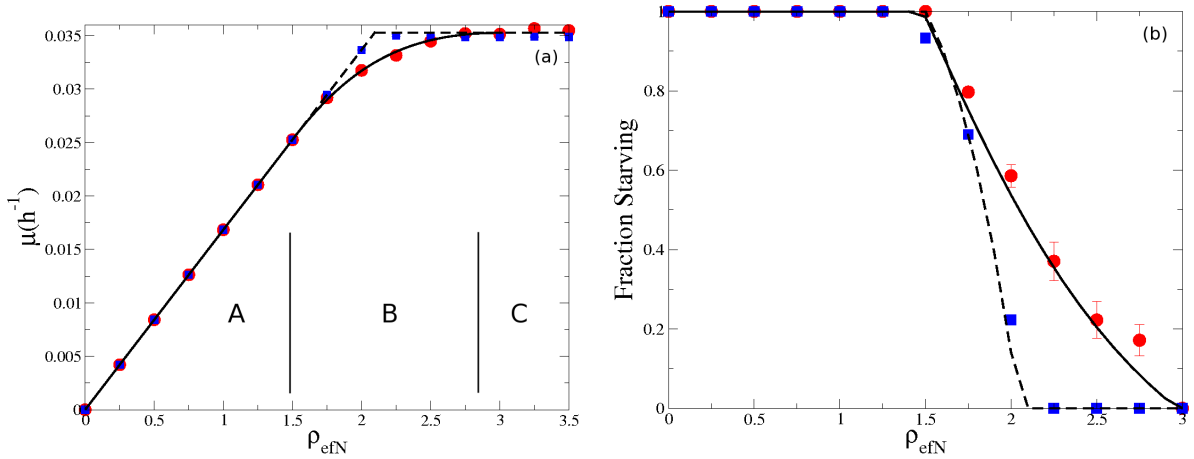
Figure 4.2: (a) growth rate exponent μ vs. external fixed nitrogen concentration ρ_{efN} in filaments with no heterocyts. Red circles indicate filaments with zero exchange between cells ($D_C = 0$), zero growth rate stochasticity ($\Delta = 0$), and high leakage (high D_M), while blue squares indicate filaments with the usual exchange ($D_C = 1.54$), growth rate stochasticity ($\Delta = 4.5$ h), and zero leakage ($D_M = 0$). Black lines are the corresponding analytical curves from Equation 4.8. "A", "B", and "C" indicate the efN ranges corresponding to the different regimes of Equation 4.8. (b) fraction of cells starving in the filament (a starving cell is defined by $N(i) = 0$) vs. external fixed nitrogen concentration ρ_{efN} for filaments with no heterocysts. Red circles and blue squares indicate the same type of filaments as in (a) and the black curves are from Equations 4.9 and 4.10.

We first examined systems with no heterocysts, no exchange between cells ($D_C = 0$), zero growth rate stochasticity ($\Delta = 0$), and high leakage (D_M high). These conditions provide a filament of isolated cells that cannot share excess fixed nitrogen and cannot retain any unused fixed nitrogen they have imported from outside the filament. We choose these conditions to be able to calculate the growth rate constant μ for the filament, which we do here.

There are two possible cases for individual cells: fixed nitrogen starvation or excess fixed nitrogen. When a cell is fixed nitrogen starved the amount imported in a small time Δt, $D_O \rho_{efN} l \Delta t$, is equal to the amount of nitrogen used for growth, $g\frac{dl}{dt}\Delta t$. This yields

$$l(t) = l_o e^{\frac{D_O \rho_{efN}}{g}t} \tag{4.2}$$

A cell is starving when $\frac{dl}{dt} = R = D_O \rho_{efN} l/g$, or equivalently when $\rho_{efN} = Rg/(D_O l)$. When a cell is importing excess fixed nitrogen it simply grows at its optimum growth rate R^{opt}

$$l(t) = l_o + R^{opt} t \tag{4.3}$$

There are three cases for what fixed nitrogen conditions a cell can experience during a cell cycle. The cell could always be starving; this occurs when $\rho_{efN} < R^{opt}g/(D_O/l_{max}) = 1.48 \times 10^{22}$ m^{-3} (assuming $R^{opt} = T_D/l_{min} = 20h/2.25\mu$m). The cell length will always follow Equation 4.2 and has doubling time

$$t_{D1} = \frac{g \cdot log(2)}{D_O \rho_{efN}} \tag{4.4}$$

The cell could be starving at the beginning of the cell cycle, when it is too short to bring in enough fixed nitrogen to grow optimally, and have available excess fixed nitrogen for the latter part of the cell cycle; this will occur when 1.48×10^{22} m$^{-3} = \frac{R^{opt}g}{D_O l_{max}} < \rho_{efN} < \frac{R^{opt}g}{D_O l_{min}} = 2.97 \cdot 10^{22}$ m^{-3} and the transition between starving and excess fN occurs when $l = l_t = R^{opt}g/D_O\rho_{efN}$. The cell will first grow at the starving rate until it reaches l_t according to Equation 4.2 over a time

$$t_1 = \frac{g}{D_O \rho_{efN}} log\left(\frac{R^{opt}g}{D_O \rho_{efN} l_{min}}\right) \tag{4.5}$$

The cell reaches l_t at t_1 and then grows until it reaches $2l_{min}$ according to Equation 4.3 over a time

$$t_2 = \frac{2l_{min}}{R^{opt}} - \frac{g}{D_O \rho_{efN}} \tag{4.6}$$

The total doubling time is

$$t_{D2} = t_1 + t_2 = \frac{g}{D_O \rho_{efN}}\left[log\left(\frac{R^{opt}g}{D_O \rho_{efN} l_{min}}\right) - 1\right] + \frac{2l_{min}}{R^{opt}} \tag{4.7}$$

The growth of cells that always experience excess fixed nitrogen, occurring when $\rho_{efN} > \frac{R^{opt}g}{D_O l_{min}} = 2.97 \times 10^{22}$ m^{-3}, follows Equation 4.3 and have doubling time $t_{D3} = l_{min}/R^{opt}$.

Growth rate constants are simply determined from the doubling time by $\mu = log(2)/t_D$.

$$\mu = \begin{cases} D_O\rho_{efN}/g & \text{if } 0 < \rho_{efN} < R^{opt}g/(D_O/l_{max}) \text{ "A"} \\ \dfrac{log(2)}{\frac{g}{D_O\rho_{efN}}\left[log\left(\frac{R^{opt}g}{D_O\rho_{efN}l_{min}}\right)-1\right]+\frac{2l_{min}}{R^{opt}}} & \text{if } R^{opt}g/(D_O/l_{max}) < \rho_{efN} < R^{opt}g/(D_O/l_{min}) \text{ "B"} \\ R^{opt} \cdot log(2)/l_{min} & \text{if } \rho_{efN} > R^{opt}g/(D_O/l_{min}) \text{ "C"} \end{cases}$$

$$(4.8)$$

This is plotted as the solid black line in Figure 4.2(a), which shows the three regimes: "A", "B", and "C". For the regime at the lowest ρ_{efN}, μ increases linearly with ρ_{efN}, as all cells are starving and growth increases with available fixed nitrogen. Once ρ_{efN} is large enough, some cells spend part of the cell cycle not starving, and then μ increases nonlinearly. When ρ_{efN} is even larger still none of the cells are starving at any point in the cell cycle. The growth levels out and does not increase with further increases in ρ_{efN}. Simulation data for this case of no exchange and no accumulation is plotted in Figure 4.2(a) as red circles.

For filaments with exchange and accumulation of fixed nitrogen, we do not follow the second, intermediate regime, and only have the first, all cells are starving regime, and the third no cells are starving regime. This is because with exchange and accumulation the filament will starve or not starve as a whole and the first regime of starvation will hold until all cells are provided for. This curve is plotted as the dashed black line in Figure 4.2(a). Simulation data for this case with exchange and accumulation of fixed nitrogen is plotted in Figure 4.2(a) as blue squares.

We determine the fraction of the cells starving by assuming that it is equal to the fraction of time spent starving. Only cells in the range $R^{opt}g/(D_O/l_{max}) < \rho_{efN} < R^{opt}g/(D_O/l_{min})$ will spend part of the cell cycle starving and the rest not starving. Cells in lower ρ_{efN} will always starve, and those in higher ρ_{efN} will never starve. The time in the cell cycle spent starving is t_1, while the total time of the cell cycle is $t_1 + t_2$, so the fraction of time spent starving is

$$F = \frac{t_1}{t_1 + t_2} \qquad (4.9)$$

This curve is the black solid curve plotted in Figure 4.2(b). This prediction of fraction starving is for filaments with no exchange between cells and no accumulation of fixed

nitrogen. Simulation results for cells in these conditions are plotted in Figure 4.2(b) as red circles.

We now take exchange and accumulation of fixed nitrogen into consideration. We assume that all non-starving cells have the same length, halfway between the length at which they stop starving, l_t, and the length at which they divide, $2l_{min}$: $l_{NS} = 0.5(l_t + 2l_{min})$. Extra fixed nitrogen will accumulate in these cells at a rate $\Delta N_{NS} = 0.5D_O\rho_{efN}(l_t + 2l_{min}) - R^{opt}g$. Similarly, the fixed nitrogen necessary to lift a starving cell out of starvation is $\Delta N_S = R^{opt}g - 0.5D_O\rho_{efN}(l_t + l_{min})$. The number of starving cells that a single non-starving cell can lift from starvation is $\Delta N_{NS}/\Delta N_S$. Multiplying this by the fraction of cells that are not starving will tell us what fraction of the total cells can be lifted from starvation by exchange of excess fixed nitrogen. The fraction of the total cells lifted from starvation is then

$$\Delta F = \frac{t_2}{t_1 + t_2}\frac{\Delta N_{NS}}{\Delta N_S}$$

The fraction starving with exchange, F_{ex}, is then F from Equation 4.9 subtract this ΔF

$$F_{ex} = \frac{t_1}{t_1 + t_2} - \frac{t_2}{t_1 + t_2}\frac{\Delta N_{NS}}{\Delta N_S} \tag{4.10}$$

This curve is the black dashed line plotted in Figure 4.2(b). Simulation results for the with exchange and accumulation of fixed nitrogen are shown in Figure 4.2(b) as blue squares.

Comparing the simulation data and curve for no exchange and no accumulation with that of exchange and accumulation in Figure 4.2(b) illustrates that exchange and accumulation allow fewer cells to starve, as is expected. It is also consistent with the higher growth rate at intermediate fixed nitrogen levels in Figure 4.2(a); less starvation results in higher growth.

4.4 Results with heterocysts

4.4.1 No leakage

With heterocysts, we first examined systems with heterocysts that were randomly placed in filaments with zero leakage. Qualitatively there are two growth regimes: starving or excess fixed nitrogen. At small heterocyst fractions and small levels of
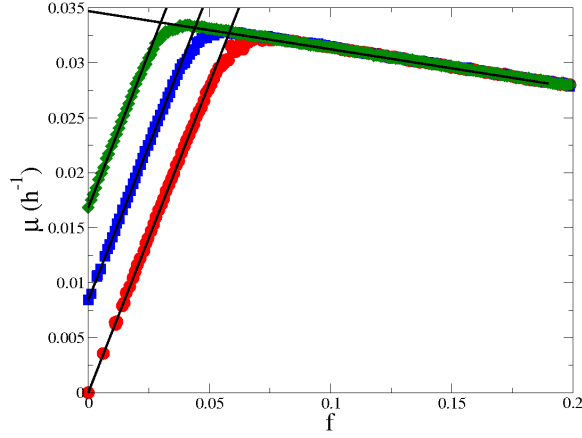
Figure 4.3: Growth rate exponent μ vs. heterocyst frequency f for systems with zero leakage ($D_M = 0$) and random heterocyst placement. Red circles indicate $\rho_{efN} = 0$, blue squares indicate $\rho_{efN} = 0.5 \times 10^{22}$ m^{-3}, and green diamonds indicate $\rho_{efN} = 1 \times 10^{22}$ m^{-3}. Black lines indicate asymptotic behavior from Equations 4.11 and 4.12.

efN, starved growth will be determined by the amount of fN produced by heterocysts as well as fN imported from outside the filament. First we set the fN consumed by growth equal to the fN produced by heterocysts and imported from outside the filament.

$$NRg = G_{het}Nf + D_O\rho_{efN}L$$

N is the total number of cells in the filament, f is the fraction of cells that are heterocysts, and L is the length of the entire filament. We can say that $NR = dL/dt$ and $N = L/l_{avg}$ where l_{avg} is the length of the average cell. This gives us

$$g\frac{dL}{dt} = G_{het}\frac{L}{l_{avg}}f + D_O\rho_{efN}L = \left(\frac{G_{get}}{l_{avg}} + D_O\rho_{efN}\right)L$$

This rearranges to

$$\frac{dL}{dt} = \left(\frac{G_{het}}{g \cdot l_{avg}}f + \frac{D_O}{g}\rho_{efN}\right)L$$

The growth constant μ is the expression in the brackets, so

$$\mu_{starve} = \frac{G_{het}}{g \cdot l_{avg}}f + \frac{D_O}{g}\rho_{efN}, \tag{4.11}$$

At large heterocyst fractions or with large levels of efN, sufficient fN is present for growth but only the vegetative cells, with fraction $1 - f$, will grow:

$$\frac{dL}{dt} = (1 - f)NR = (1 - f)\frac{L}{l_{avg}}R = (1 - f)\frac{R}{l_{avg}}L$$

$l_{min} = l_{avg} \cdot ln(2)$, and the coefficient of L is μ so

$$\mu_{excess} = \frac{R \cdot ln(2)}{l_{min}}(1 - f), \qquad (4.12)$$

Both of these limiting behaviors μ_{starve} and μ_{excess} are shown as straight black lines in Figures 4.3.

The modelling results for filaments with different levels of efN and with zero leakage ($D_M = 0$), shown in Figure 4.3, agree well with the limiting behaviours for small or large f. There is a clear optimal heterocyst frequency, f^*, at which growth is maximal. Increasing the amount of efN shifts μ_{starve} up, while μ_{excess} is unaffected. As efN increases, the optimal frequency decreases and the corresponding maximal growth rate μ^* increases.
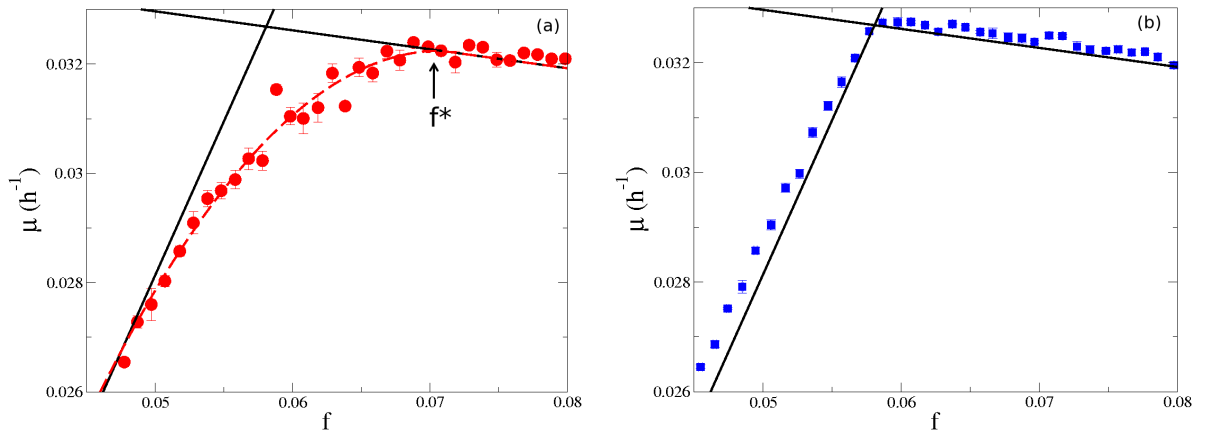
Figure 4.4: (a) highlights the data near the maximum μ in Figure 4.3. The red dashed curve indicates the best fit near the indicated optimal heterocyst fraction f^*. (b) data near the maximum μ in Figure 4.3 except with infinite diffusivity.

The modelled growth rates are somewhat below the limiting regimes near f^*, as highlighted in Figure 4.4(a). This growth deficit appears to be due to finite diffusivity of fN within the filament. With heterocysts placed randomly in the filament, near f^* some sections of the filament have more heterocysts than necessary to fulfill the fixed nitrogen needs of all vegetative cells and other sections have fewer heterocysts than necessary to achieve this. With no leakage from the filament, one would expect that the excess fixed nitrogen from the sections with extra heterocysts would diffuse to the sections with too few heterocysts. This is not the case because the filament is growing, and on large enough scales growth is able to outpace the diffusion.

We tested this hypothesis by implementing infinite diffusivity by taking $D_C \rightarrow \infty$, expecting that this would then give agreement between the numerical data and the analytical lines and yield a clear distinction and transition between filaments that are starving and filaments that have excess fixed nitrogen. This is plotted in Figure 4.4(b), showing two linear regions and sharp intersection, but not agreement between analytical and numerical. The numerical data is consistently higher than the analytical prediction in the starving regime. $\mu(f)$ for the starving regime is predicted by Equation 4.11. The only part of this equation that could change is l_{avg}, and the higher μ implies that infinite diffusivity results in a smaller l_{avg}.

Infinite diffusivity means that all cells have the same concentration of fixed nitrogen, implying that larger cells have more fixed nitrogen than smaller cells. In a starving filament, this means that larger cells are able to grow more than smaller cells, so that cells are small for a larger fraction of their time to division and are big for a smaller fraction of their time to division. This explains a smaller l_{avg}. A smaller l_{avg} also affects the number of heterocysts. In a filament with shorter cells, there are more cells per length of filament, and therefore more heterocysts at a given heterocyst frequency. In our model, heterocysts produce the same amount of fixed nitrogen regardless of their length, and so more heterocysts in the filament results in more fixed nitrogen provided to the filament. This increase in fixed nitrogen production is the source of the increased growth rate.

The growth deficit seen in Figure 4.4(a) is smaller when heterocysts are placed close to starving cells (see below). We do not have an analytical treatment of this growth deficit, though it is intriguing. As a result of the growth deficit, the optimal heterocyst frequency from our quantitative model, f^*, is larger than given by the intersection of μ_{starve} and μ_{excess}, as shown in Figure 4.5(a).

The maximum growth rates, μ^*, corresponding to the optimal heterocyst frequencies f^*, are plotted with a solid red line vs. ρ_{efN} in Figure 4.5(b). The dashed orange line shows the growth rate with no heterocysts ($f = 0$), which exhibits sharply reduced growth at smaller ρ_{efN}. The dashed black line shows the growth rate exhibited by the fixed heterocyst fraction f_0 that is optimal for $\rho_{efN} = 0$, which exhibits a constant but reduced growth when $\rho_{efN} > 0$. In general, filaments that maintain an optimal heterocyst fraction by maintaining $f = f^*$ as ρ_{efN} varies will outgrow
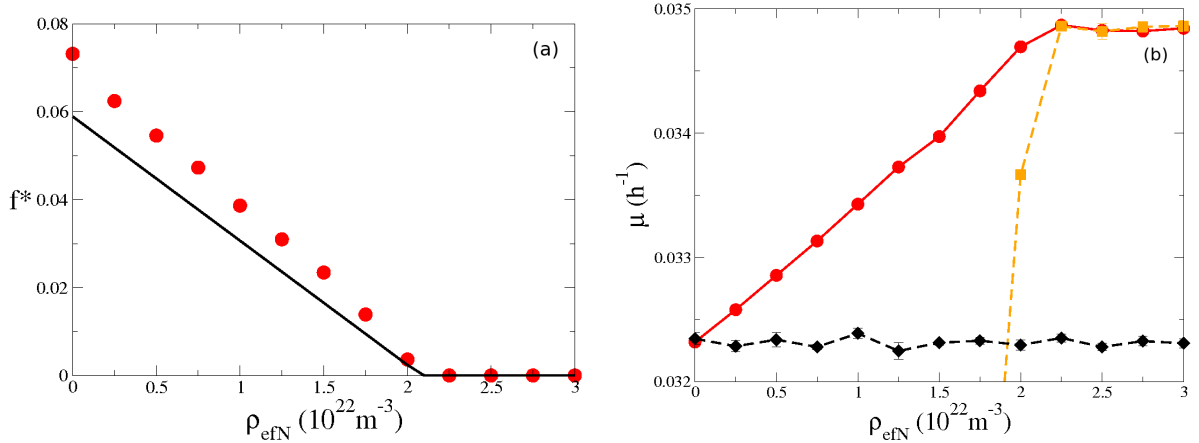
Figure 4.5: (a) optimum heterocyst frequency f^* vs. ρ_{efN} using random heterocyst placement with zero leakage. Numerical data is shown as red circles while the frequency described by the intersection of Equations 4.11 and 4.12 is shown as a solid black line. Disagreement between the red circles and black line is due to the growth deficit highlighted in Figure 4.4(a). (b) growth rate μ vs. ρ_{efN}. Red circles with a solid line indicate μ^* from random heterocyst placement with zero leakage. Orange squares with a dashed orange line indicate μ for a filament with no heterocysts, while the black diamonds and black dashed line indicate μ for a filament maintaining a constant heterocyst frequency that is optimal at $\rho_{efN} = 0$.

filaments with any given fixed heterocyst fraction.

4.4.2 Leakage

As shown in Figure 4.6, with no leakage the random (solid red line) and no-Mch (dashed red line) heterocyst placement strategies have almost indistinguishable growth rates. Local heterocyst placement (red circles) leads to slightly faster growth with short wait times $\tau = 1$ h, but shows significantly slower growth at longer wait times comparable to heterocyst maturation times. Note that f is not independently controlled with local placement, so that only a narrow range of f is seen as τ is varied. Faster local growth is seen with smaller τ as starving cells are provided with fN earlier, and experience shorter periods of restricted growth.

With non-zero leakage, the three heterocyst placement strategies produce noticeably different growth curves. At 1% leakage (blue lines and squares), the growth rates with no-Mch are slightly above the random placement strategy. More striking is the dramatic improvement with local placement, which has significantly better growth at
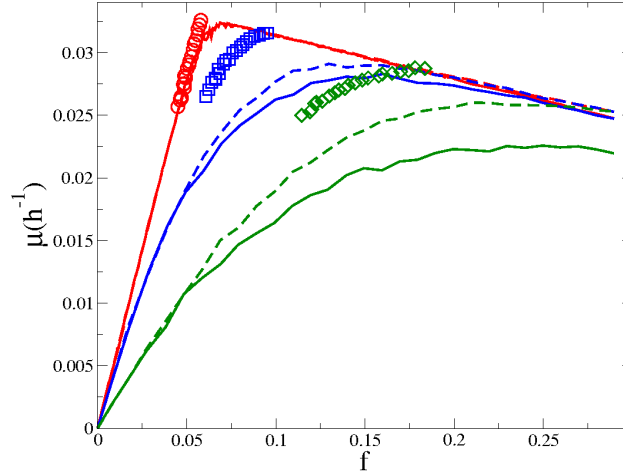
Figure 4.6: Growth rate constant μ vs. heterocyst frequency f for different heterocyst placement strategies and leakage levels with zero external fixed nitrogen. Zero leakage is shown in red, 1% leakage is shown in blue, and 10% leakage is shown in green. Solid lines are random heterocyst placement, dashed lines are no-Mch heterocyst placement. Red circles, blue squares, and green diamonds show data for local heterocyst placement for zero, 1%, and 10% leakage respectively. The local heterocyst placement data points show variation of the period of starvation until commitment τ from 1-20 h, with shorter τ exhibiting larger μ.

any given heterocyst frequency, but also better maximal growth μ^* for $\tau \in [1, 13]$h. The same trends continue when leakage is increased to 10% (green lines and green diamonds): the growth rates of filaments with local heterocyst placement are higher than the maximal growth rates of other strategies for $\tau \in [1, 17]$h, and we also see that the corresponding heterocyst frequencies f^* are much lower with local strategies than with either random placement or no-Mch placement strategies.

Note that f^* and μ^* correspond to the heterocyst frequency and growth rate exhibited by the local placement strategy for a given delay time τ, or to the *growth-optimized* heterocyst frequency and corresponding growth rate exhibited by the random or no-Mch placement strategies.

In general, leakage of fN from the cyanobacterial filament will more strongly inhibit growth if heterocysts are not placed close to starving vegetative cells. In any case, leakage decreases the growth rate at a given heterocyst frequency, and so leads to a larger heterocyst frequency f^* and correspondingly decreased growth rate μ^* with both the local strategy and the growth-optimized random strategies.

4.4.3 Varying external fixed nitrogen concentration
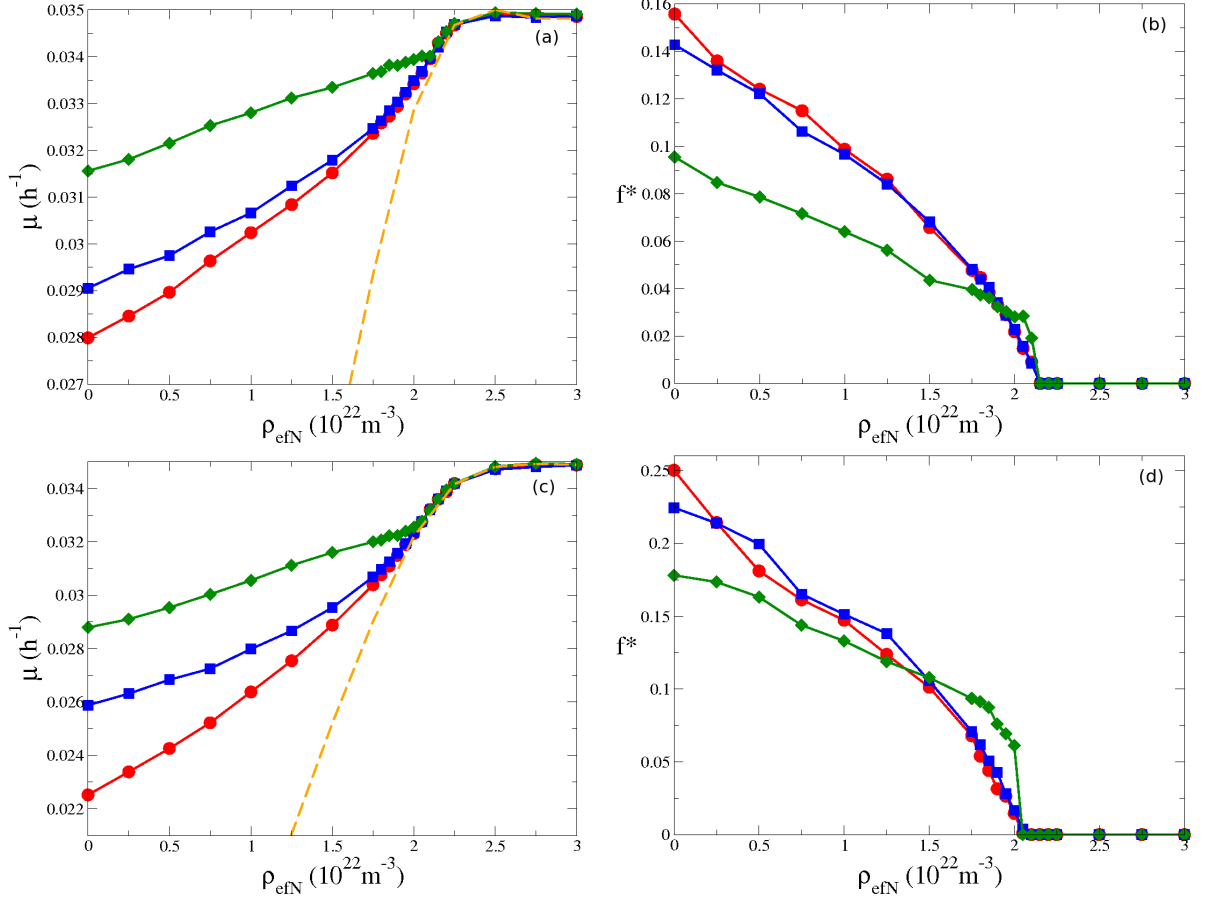
Figure 4.7: (a) growth rate constant μ vs. external fixed nitrogen concentration ρ_{efN} for filaments with 1% leakage. Red circles with a solid line are random heterocyst placement at the optimum frequency, blue squares with a solid line are no-Mch heterocyst placement at the optimum frequency, green diamonds with a solid line are local heterocyst placement, and the orange dotted line is filaments with no heterocysts. (b) heterocyst frequency at which maximum growth occurs vs. external fixed nitrogen concentration ρ_{efN}. Red circles with solid line are random heterocyst placement, blue squares with solid line are no-Mch heterocyst placement and green diamonds with solid line are local heterocyst placement. (c) and (d) are similar runs labelled in the same fashion as (a) and (b), respectively, with higher leakage of 10%.

When the external fixed nitrogen concentration, ρ_{efN}, is increased from zero the differences between the heterocyst placement strategies diminish. Figure 4.7(a) shows growth, μ^*, vs. ρ_{efN} for all three placement strategies with 1% leakage. Also shown with the orange dashed line is the growth expected with no heterocysts ($f = 0$).

We see that local heterocyst placement has a significantly higher growth rate than optimized random or no-Mch strategies. The differences are largest at $\rho_{efN} = 0$, decrease as efN increases, and vanish when all growth is supported by efN alone. Similar qualitative behavior is seen with 10% leakage, in Figure 4.7(c), though the growth rates are somewhat lower.

Figures 4.7(b) and (d) show the heterocyst frequency, f^*, for all three heterocyst placement strategies for a range of efN, for 1% and 10% leakage respectively. For most of the lower fixed nitrogen levels, local heterocyst placement has a significantly lower heterocyst frequency than optimized random or no-Mch heterocyst placement. When heterocysts are placed near starving cells fewer heterocysts are necessary to satisfy the fixed nitrogen requirements for the growing vegetative cells. As f^* approaches zero, the trend reverses. The local heterocyst placement has a higher f^* than either random or no-Mch placement because local can continue to beneficially place heterocysts near starving cells as f increases and fewer cells starve, while random strategies cannot. The reversal of this trend is particularly apparent with larger leakage levels (Figure 4.7(d)), where placement of heterocysts close to fast growing vegetative cells is particularly important, and leads to a noticeable nonlinearity of the heterocyst frequency f^* with local placement at larger ρ_{efN}.

4.4.4 Heterocyst spacing

Figures 4.8(a)-(c) show heterocyst spacing distributions for the different heterocyst placement strategies with 1% leakage, all with approximately $f = 0.1$ for ease of comparison. Both random placement and no-Mch favour small heterocyst spacings. Random placement, in Figure 4.8(a), peaks at adjacent heterocysts (corresponding to the Mch phenotype) and drops off for larger separations, while no-Mch placement, in Figure 4.8(b), peaks at a spacing of 4 intercalating cells between heterocysts. Significant bias towards even spacings, due to ongoing filament growth, is also seen. The distribution for local placement is quite different, with a symmetric peak at approximately 12 intercalating cells and very few heterocysts separated by less than 6 cells. Figure 4.8(d) is an experimental spacing distribution after 96h of fixed nitrogen deprivation for WT *Anabaena* PCC 7120 [26]. The heterocyst spacing distribution with the local placement strategy is qualitatively similar to the experimental distribution,

even though the strategy selection was done with respect to growth alone — without consideration of the spacing distribution.
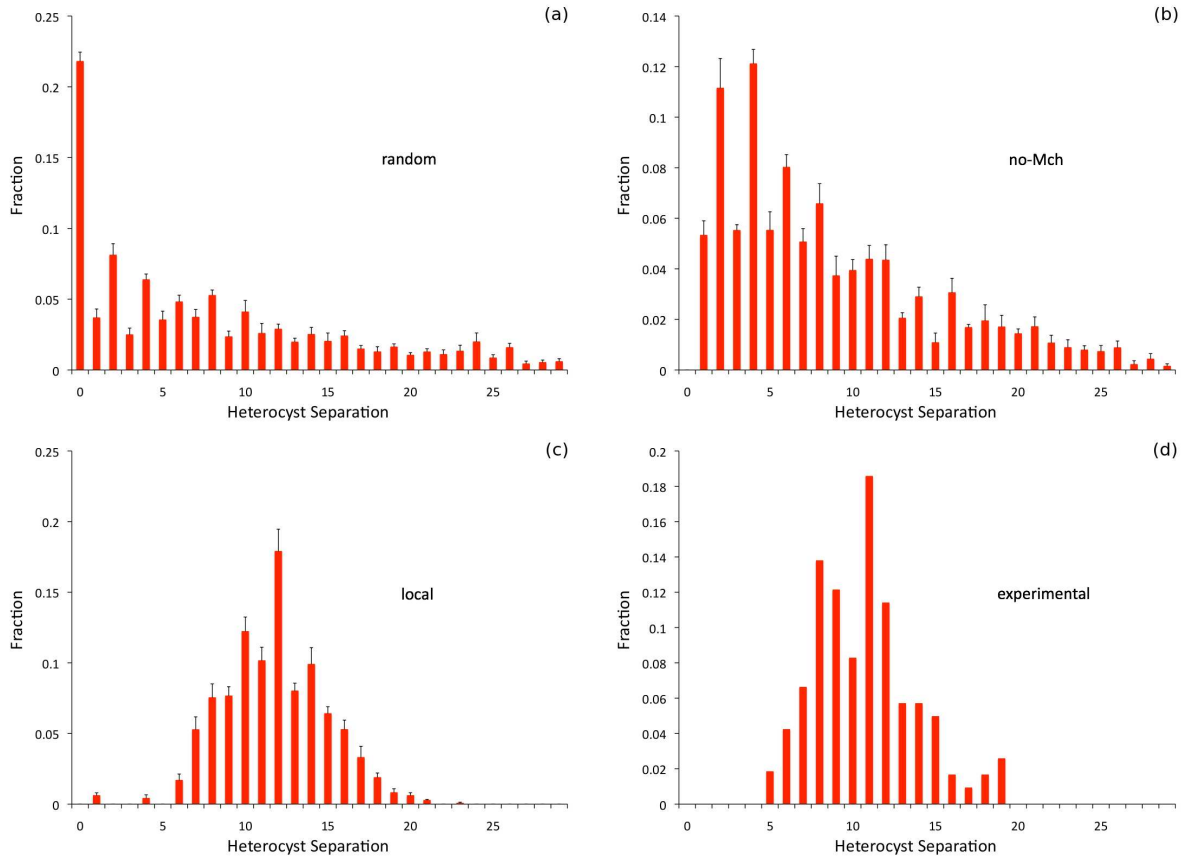
Figure 4.8: (a)-(c) are heterocyst spacing distributions from stochastic simulations with 1% leakage and zero external fixed nitrogen with error bars representing the standard deviation of the mean. (a) is random heterocyst placement with 10% heterocysts, (b) is no-Mch heterocyst placement with 10% heterocysts, and (c) is local heterocyst placement with period of starvation until commitment of $\tau = 8$h after 5 days of growth. (d) is the experimental steady state WT distribution from [26] after 96 hours.

With non-zero ρ_{efN} we consider only the heterocyst spacing distributions for the local heterocyst placement strategy, as illustrated in Figures 4.9(a)-(c) with 1% leakage. The peak separation increases with ρ_{efN}, consistent with the decreasing f^* we saw in Figure 4.7 (b). The distribution also becomes significantly wider. By scaling the heterocyst separation by the average separation for each distribution we see in Figure 4.9 (d) that the distributions at different ρ_{efN} approximately collapse to a single scaled distribution independent of ρ_{efN}. The inset in Figure 4.9(d) is the average

heterocyst separation vs. ρ_{efN}.
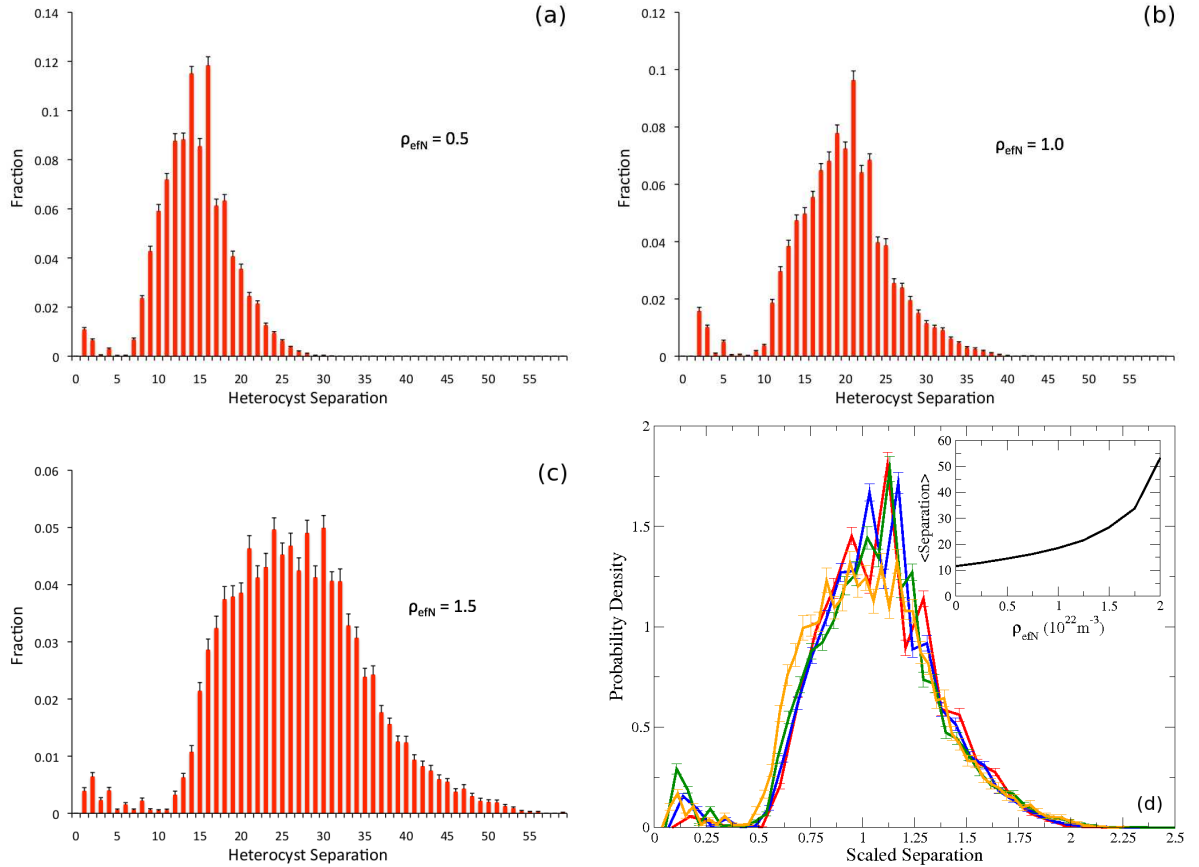
Figure 4.9: Heterocyst spacing distributions for filaments with local heterocyst placement, 1% leakage for various ρ_{efN} and a period of starvation until committment of $\tau = 8$h. (a) $\rho_{efN} = 0.5 \times 10^{22}\text{m}^{-3}$, (b) $\rho_{efN} = 1 \times 10^{22}\text{m}^{-3}$, and (c) $\rho_{efN} = 1.5 \times 10^{22}\text{m}^{-3}$. (d) shows the probability density of the distributions with $\rho_{efN} = 0$ (red), 0.5 (blue), 1 (green), and 1.5 (orange) $\times 10^{22}\text{m}^{-3}$ vs. scaled heterocyst separation. Inset in (d) is the average heterocyst separation vs. efN concentration.

4.5 Discussion

We first explored the growth and starvation of filaments with no heterocysts with a quantitative model for fixed nitrogen (fN) transport and dynamics that includes growth of vegetative cells. Filaments in low fixed nitrogen increase their growth rate linearly with the external fixed nitrogen level until their growth needs are satisfied, when the growth rate levels off for all higher external fixed nitrogen levels. For lower external fixed nitrogen, all cells starved. For intermediate levels, some cells starved

and once the external fixed nitrogen level was high enough, no cells starved. The end of intermediate range coincided with the concentration at which the growth rate plateaued.

We then explored three heterocyst placement strategies in a cyanobacterial filament using the same quantitative model with production of fN by heterocysts. The three strategies were random placement, random placement with no-Mch, and local placement. For random strategies, we found an optimal heterocyst frequency f^* at which the growth rate of the filament is maximized, μ^*. With fewer heterocysts than f^*, filaments are fN limited. With more heterocysts than f^*, the filament is hampered by an excess of non-growing heterocysts. As the external fN, ρ_{efN}, is increased from zero, the optimal heterocyst frequency decreases continuously until it reaches zero at a concentration ρ^*_{efN} where all fN needs are met by imported extracellular fN. The local placement strategy led to a similar decreasing heterocyst fraction f^* with ρ_{efN}, though without explicit optimization of the heterocyst frequency.

Without leakage of cytoplasmic fN from the filament, via D_M, the difference between these strategies is small (see Figure 4.6, red lines and circles). With leakage there are significant growth differences between the strategies under limiting efN conditions (with $\rho_{efN} < \rho^*_{efN}$). Filaments using local placement of heterocysts grow faster than filaments with no-Mch, which in turn grow faster than filaments using random heterocyst placement. The differences between the strategies are largest when $\rho_{efN} = 0$, and decrease as ρ_{efN} increases, until it vanishes above ρ^*_{efN} where the heterocyst frequency $f^* = 0$ for all three strategies. We suggest that leakage of fN from growing filaments may be important for understanding the adaptive nature of heterocyst placement strategies. Moving forward we will use 1% leakage, because Figure 4.7(b) indicates that at $\rho_{efN} = 0$ the heterocyst frequency is close to 10% for the local strategy, which is about right for PCC 7120, while Figure 4.7(d) indicates that at $\rho_{efN} = 0$ the heterocyst frequency is about 18%, which is too high for PCC 7120.

We believe that the observed growth differences are relevant. Selective sweeps occur when a beneficial mutation 'sweeps' a population and becomes fixed [120]. Selective sweeps have been observed in cultures of E. coli for mutations with a fitness parameter as small as $m = 0.006$ [120], which in our system corresponds to a growth

rate constant difference of 0.0003h^{-1}. Figure 4.7(a) shows that the growth rate difference between random heterocyst placement strategies and local heterocyst placement is about 0.0025h^{-1} for 1% leakage for zero external fixed nitrogen, and is similar for 10% leakage — substantially larger than necessary for a selective sweep in *E. coli*.

Significantly, we did not pick our best strategy to match the heterocyst spacing pattern nor did we tweak the heterocyst frequency f. Rather we implemented a simple local strategy which itself chose f^*. This best, local strategy resulted in faster growth than optimized random strategies, and in heterocyst spacings remarkably similar to those seen experimentally. The similarity of our model heterocyst spacing distributions with observed patterns suggests that local fN starvation may drive heterocyst development. Indeed, during steady state development we believe that the extensive genetic network of *Anabaena* heterocyst development [84] effectively implements something similar to our local strategy. This is consistent with experimental work [8] (see also [75]) showing a dip in the fixed nitrogen level approximately halfway between two widely separated heterocysts.

We note that the heterocyst spacing patterns themselves are not adaptive. Rather the patterns reflect an adaptive local strategy for placing heterocysts close to locally starving cells. This starvation is a combination of distance from heterocysts and fast growing cells. Local placement is particularly adaptive when fN leaks from the cyanobacterial filament (see [2, 60, 88, 93, 94, 107], as discussed in Section 4.1).

Qualitatively our best strategy is easy to state: "quickly differentiate cells that are locally starving of fixed nitrogen". Effectively implementing the strategy is not trivial. We note two complications. The first is that our local strategy works for ongoing differentiation of heterocysts, during steady state growth. In that regime, long after efN deprivation, a dedicated mechanism to avoid Mch is not needed because nearby cells begin to starve at different times. However, in the first burst of differentiation we need to avoid Mch with a temporary mechanism to prevent differentiation of the entire filament. Indeed, the initial heterocyst spacing is distinct from the steady state pattern [26] and cyanobacterial filaments are thought to use diffusible inhibitors derived from PatS to suppress Mch [6]. The second is that we have simplified heterocyst commitment [26] and subsequent delay before *de novo* nitrogen fixation into a single delay τ. While we find that a smaller τ always leads to

faster filament growth, heterocysts do take significant time to both commit and to begin to fix nitrogen [26, 27, 28, 84] (though see [56]). While we have explored a correspondingly large range of $\tau \in \{1-20\}$h, with similar results throughout, the details may be expected to change with a more detailed model of heterocyst commitment. As such, a more detailed model of heterocyst commitment is explored in Chapter 5.

We predict a plastic developmental response of heterocyst frequency to levels of efN, where the heterocyst frequency f decreases rapidly with ρ_{efN}. This implicitly assumes that there is no threshold extracellular fixed nitrogen concentration above which heterocysts will not differentiate, and more broadly that there is not a fixed developmental pattern of heterocysts. Supporting a plastic developmental response, early work by Fogg [2] showed a time dependent heterocyst frequency that increased with decreasing levels of efN, Ogawa and Carr [118] found heterocyst frequency and nitrate concentration to be inversely related, and Horne *et al* [119] found the same correlation in field studies. Fogg also observed significant heterocyst development at 40 μM extracellular fixed nitrogen (Figure 2 of [2]), though not in steady state conditions. Plastic developmental responses are observed in systems ranging from plants [121] to the brain [122]. We feel that filamentous cyanobacteria are particularly amenable to exploring and understanding the adaptiveness of this plastic response. The time is now ripe to undertake constant external fixed nitrogen concentrations with single filament studies in the chemostat-like environment of microfluidic systems, such as filamentous cyanobacteria within the long channels used in studies of persister cells of *E. coli* [123, 124].

Our heterocyst placement strategy is deliberately simple and designed to be sufficient to investigate steady state heterocyst patterns. Chapter 5 discusses a more detailed but still coarse grained model for heterocyst differentiation. The more detailed model approximates commitment as the time at which fixed nitrogen storage (also introduced in Chapter 5) runs out, and explicitly includes the effects of the gene products of *patS* and *hetN* in forming an inhibition range around committed and mature heterocyst cells. In contrast to this chapter, the next chapter also separates the timing of commitment and fixed nitrogen production.

Our model and its results assume that, apart from fixed nitrogen, there are no nutrient or metabolite requirements that limit growth differentially along the filament.

Doubtless, there are experimental conditions where this assumption is not warranted. Certainly, there are interesting nutrients that we have not included in our model. For example, the supply of carbohydrates from vegetative cells to heterocysts [125, 126] was not considered. Nevertheless, while carbohydrate transport would be expected to starve the interior cells of clusters of contiguous heterocysts and reduce their capacity for nitrogen fixation, our primary focus was on the local heterocyst placement strategy under conditions of nonzero external fixed nitrogen, where no clusters of heterocysts are observed and heterocysts are broadly spaced. However, carbohydrate limitation would affect random placement strategies, which might then have even more reduced growth compared to our best, local strategy.

For cyanobacterial filaments implementing a local heterocyst placement strategy we have several experimentally testable predictions that we expect to be observed in steady state conditions, which are independent of the specific parameterization of our quantitative model. To explore them in the lab, external concentrations of fixed nitrogen (ρ_{efN}) would need to be controlled by microfluidic devices [123, 124] or flow cells. The first is that heterocyst frequency f^* will decrease rapidly and continuously as ρ_{efN} is increased from zero. As shown in Figures 4.7 (b) and (d), the dependence is markedly nonlinear with larger leakage rates of fN from the growing filament. The second, related, prediction is that the peak of the heterocyst spacing distribution will increase as ρ_{efN} increases. This is illustrated in Figures 4.9 (a)-(c). We also expect that the width, or standard deviation, of the heterocyst spacing will increase as the mean spacing increases. Indeed, for our model, the standard deviation and mean spacing are approximately proportional.

Our results may also have implications for quantitative models of biogeochemical cycling of fixed nitrogen in marine and lake environments, where filamentous cyanobacteria can play a significant role. We note that a number of existing models [117, 127, 128] use growth rates and nitrogen fixation rates that are independent of the biologically available fixed nitrogen. We find that, with a local heterocyst placement strategy, the filament growth rate μ^* is approximately (but not exactly) independent of ρ_{efN} — see Figure 4.7 (a) and (c), consistent with the approximately constant growth reported for chemostat studies by Elder and Parker [29]. However we expect that *de novo* nitrogen-fixation to be proportional to f^*, which in turn strongly

decreases with increasing ρ_{efN}.

Chapter 5

De Novo Heterocyst Patterning

5.1   Motivation and background

Chapter 3 looked at the nitrogen distribution in the filament, while Chapter 4 examined the growth and heterocyst patterns in the filament during steady state growth. In this chapter we want to investigate the initial pattern formation of heterocyst differentiation following nitrogen deprivation. This process was discussed in Section 2.4 but I will briefly summarize here.

Upon nitrogen deprivation, the gene *ntcA* is activated [49], which activates the gene *hetR* [19], which is regarded as the master switch of heterocyst differentiation [25]. Both individual cells and clusters of cells in the cyanobacterial filament begin differentiation into heterocysts. HetR turns on the gene *patS* [56], which inhibits differentiation [6, 26] and participates in the resolution of differentiating clusters into single cells that complete differentiation, while the other cells of each cluster revert to a vegetative state [26]. In filaments with mature heterocysts, *hetN* maintains the heterocyst pattern [61, 62].

When deprived of fixed nitrogen, the filament growth slows [53]. In *Anabaena* PCC 7120, cells commit to heterocyst differentiation between approximately 8 and 14 hours after nitrogen deprivation [26]. After about 18-24 hours the heterocysts have matured and begin to fix nitrogen [42, 59], and occur in a regular pattern in the filament.

It is not clear how committing cells are selected from the cells that have begun differentiation, although it is thought to be related to *hetR* and *patS* expression. It has been suggested [69] that the uneven accumulation of nitrogen storage would cause the cells with the lowest storage to notice nitrogen starvation first.

The heterocyst differentiation model of Chapter 4 did a good job of reproducing the experimental steady state heterocyst pattern. However, the model relies on a lack of cytoplasmic fixed nitrogen and a long waiting time, and this is not sufficient to

model the initial heterocyst differentiation upon nitrogen deprivation. In steady state, existing heterocysts provide most cells with fixed nitrogen, and only those furthest from the heterocysts will starve. In contrast, upon initial nitrogen deprivation all the cells will starve nearly simultaneously. A starvation based model in which all cells begin differentiation upon starvation will lead to most or all of the cells completing differentiation, because there are no existing heterocysts to provide fixed nitrogen to prevent differentiation. A successful model requires a means to break the starvation symmetry between the cells, and a way to prevent other cells from differentiating.

In this chapter, we have modelled the differentiation of heterocysts using the level of fixed nitrogen storage to determine which cells will commit to differentiation and break the starvation symmetry. Also in the model, when a cell commits, it inhibits nearby cells from commitment with PatS. Strong evidence exists for both fixed nitrogen storage and lateral inhibition. Cyanobacteria have fixed nitrogen storage (discussed in Section 2.5), primarily composed of cyanophycin [67] and phycobiliprotein [69, 67]. Upon nitrogen deprivation, proheterocysts produce PatS, which has been shown to resolve clusters of proheterocysts into single cells that complete differentiation [6, 26]. Using this commitment model, based on storage and lateral inhibition, we generate the heterocyst pattern at 24 hours and later times, as well as filament length, cell commitment, and heterocyst frequency in time. The pattern and other characteristics qualitatively match existing experimental measurements.

It has been suggested that heterocyst differentiation is related to the cell cycle [31, 72, 73, 129] and that only recently divided cells can become heterocysts. However, experiments have shown that cells that have not recently divided can become heterocysts [46, 53]. Despite the lack of any cell cycle criteria in our commitment model, we show that our model favours shorter cells for heterocyst differentiation.

The genes $patS$ and $hetN$ are involved in formation and maintenance of the heterocyst pattern, and $\Delta patS$ [26] and $\Delta hetN$ [61] mutants deviate significantly from the wild type heterocyst spacing pattern. We are able to remove the elements of our model corresponding to the action of PatS and HetN and qualitatively reproduce the effect of the mutants.

5.2 Model

Our model is an improvement upon the model of Chapter 4, with 1% leakage. In addition to the fixed nitrogen dynamics and the growth model built upon them, here we include the dynamics of fixed nitrogen storage. We also outline a more complex heterocyst commitment and differentiation model that depends on the amount of remaining nitrogen storage and has inhibition of the commitment of nearby cells.

5.2.1 Fixed nitrogen storage

The first new element of the model is the storage of fixed nitrogen. Stored fixed nitrogen, N_S, will provide fixed nitrogen for the cell to grow when the cytoplasmic fixed nitrogen has run out. A fraction of the fixed nitrogen incorporated by the cell, f_s, will be stored when growing on cytoplasmic fixed nitrogen and available for growth when the growth rate falls below a fraction f_g of the optimal growth rate R^{opt}.

$$\dot{N}_S = \begin{cases} f_s g R & \text{if } N_C(i,t) > 0 \text{ or } N_C(i,t) = 0 \text{ and } \Phi_{in}/g > f_g R^{opt} \\ f_s \Phi_{in} - g f_g R^{opt} & \text{if } N_C(i,t) = 0 \text{ and } \Phi_{in}/g \le f_g R^{opt} \end{cases} \quad (5.1)$$

The fraction of the optimal growth rate supported by storage, f_g, is 0.25 to be consistent with the filament length growth rate after nitrogen deprivation [53], as discussed in Section 2.4.

To determine the fraction of the fixed nitrogen that is storage, f_s, we look at the amounts of the fixed nitrogen reserves cyanophycin and phycobiliprotein in cyanobacteria.

PCC 7120 has 636 μg arginine/mg cholorophyll, while a cyanobacterial species that does not synthesize cyanophycin has 68 μg arginine/mg chlorophyll, so we estimate that 568 μg arginine/mg chlorophyll is from cyanophycin [65]. Cyanophycin is 1:1 arginine:aspartic acid, with arginine with four nitrogen atoms and aspartic acid with one nitrogen atom, so there are approximately five nitrogen atoms in cyanophycin for every arginine from cyanophycin. *Anabaena cylindrica* has 0.58 μg chlorophyll/10^6 cells, but *A. cylindrica* is 2.25 times as large as PCC 7120 [75], so we estimate that PCC 7120 has 0.26 μg chlorophyll/10^6 cells. Together, the 568 arginine/mg chlorophyll $\times$ 5 N atoms per arginine $\times$ 0.26 μg chlorophyll/10^6 cells gives 2.55$\times 10^9$ nitrogen

atoms in cyanophycin per cell. There are 2.07×10^{10} nitrogen atoms per cell in PCC 7120 (see Section 3.2.2) and so 12.3% of the nitrogen atoms are in cyanophycin. This ballpark estimate for the amount of cyanophycin from *A. cylindrica* approximately matches Li *et al* [64] who observed a maximum accumulation of cyanophycin as 20% of protein, so we choose a value of 15% of the fixed nitrogen in the cell as cyanophycin.

Phycobiliprotein can comprise up to 60% of the soluble protein in a cell [130] and upon nitrogen deprivation is observed to degrade incompletely and transiently in cells that remain vegetative, and completely in cells that become heterocysts [70]. We choose another 15% of the total fixed nitrogen in the cell to be phycobiliprotein available for growth upon nitrogen starvation. This gives a total of 30% of the fixed nitrogen as storage available for growth, and we assign $f_s = 0.3$. The dependence of our results on this value is discussed in Section 5.4.

5.2.2 Cell growth and division

The growth of a cell proceeds as in Section 4.2.2 unless the growth rate would fall below $f_g R^{opt}$. When the growth due to cytoplasmic fixed nitrogen would fall below $f_g R^{opt}$ then the storage supports growth at that rate unless storage runs out.

$$
R = \begin{cases}
R^{opt} & \text{if } N_C(i,t) > 0 \\
min\left(\Phi_{in}/g, R_i^{opt}\right) & \text{if } N_C(i,t) = 0 \text{ and } \Phi_{in}/g > f_g R^{opt} \\
f_g R^{opt} & \text{if } N_C(i,t) = 0, \Phi_{in}/g \leq f_g R^{opt}, \text{ and } N_S(i,t) > 0 \\
\Phi_{in}/g + N_S/g & \text{if } N_C(i,t) = 0, \Phi_{in}/g + N_S/g < f_g R^{opt}
\end{cases}
\tag{5.2}
$$

The first case is when there is sufficient cytoplasmic fixed nitrogen to fully support the optimal growth of the cell. The second case describes when there is not sufficient cytoplasmic fixed nitrogen to fully support the optimal growth of the cell, but the growth rate has not dropped low enough for the storage to support any growth, and so the cell grows less than optimally using the available cytoplasmic fixed nitrogen. The third case is when the growth supported by the cytoplasmic fixed nitrogen has dropped low enough for storage to provide the fixed nitrogen to maintain the growth rate at a fraction of the optimal growth rate, $f_g R^{opt}$. The final case is when there is not enough storage to support growth at $f_g R^{opt}$ and so the cell grows using whatever cytoplasmic and stored fixed nitrogen is available.

Cells divide as in Section 3.2.2, except that when $N_C = 0$ then a cell is not allowed to divide. This is consistent with the increase in filament length but not cell number at early times after fixed nitrogen deprivation in Asai *et al* [53] (see inset in Figure 5.2(a)).

The change in the cytoplasmic fixed nitrogen is similar to that described in Section 4.2.2.

$$G_{veg} = \begin{cases} -gR & \text{if } N_C(i,t) > 0 \text{ or } N_C(i,t) = 0 \text{ and } \Phi_{in}/g > f_g R^{opt} \\ -\Phi_{in} & \text{if } N_C(i,t) = 0 \text{ and } \Phi_{in}/g \leq f_g R^{opt} \end{cases} \tag{5.3}$$

5.2.3 Heterocysts
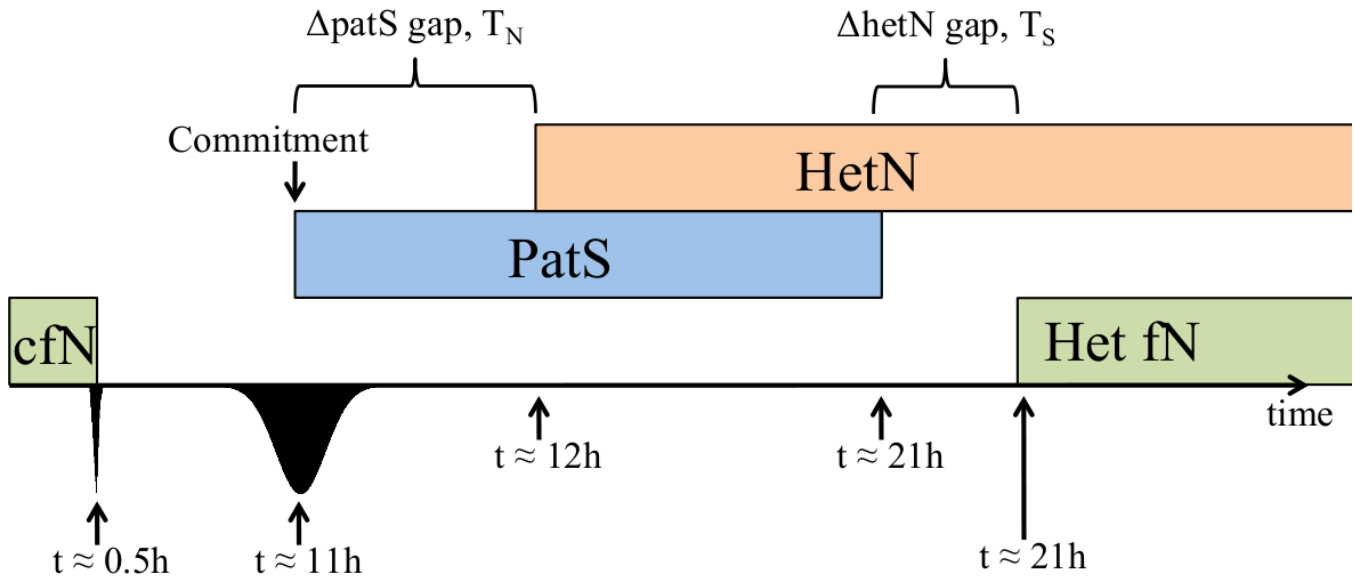
Figure 5.1: Heterocyst commitment and differentiation model. Indicated times are typical but can depend on timing parameters such as T_N and T_S. Times are not to scale. $t = 0$ when $efN \to 0$. "cfN" represents cytoplasmic fixed nitrogen, "PatS" represents the time period that PatS type inhibition is in effect, "HetN" represents the time period that HetN type inhibition is in effect, and "Het fN" represents the fixed nitrogen produced by heterocysts. We have designed this model around the existing knowledge of storage [67], PatS inhibition [6, 26], HetN inhibition [63, 61, 62], differentiation timing [42, 59, 26], and filament growth [53].

The role of storage in commitment

Our model relies on fixed nitrogen storage to determine which cells will commit to heterocyst differentiation. When cyanobacteria are deprived of fixed nitrogen, cyanophycin degrades before phycobiliprotein [67]. When the cyanophycin is depleted the cell will begin to use phycobiliprotein as a fixed nitrogen reserve. Given that heterocysts completely deplete their phycobiliprotein and vegetative cells do so incompletely and transiently, the beginning of phycobiliprotein degradation could serve as a differentiation signal. Our model commits cells to heterocyst differentiation when half or more (dependence of the results on this value are discussed in Section 5.4) of their potential storage is emptied unless they are inhibited from differentiation by a committed or mature heterocyst. Committed heterocysts do not grow: their $R^{opt} = 0$.

Lateral inhibition

Inhibition of heterocyst differentiation is performed by the products of the genes *patS* and *hetN*. Accordingly, our model includes two types of lateral inhibition, each with characteristics corresponding to what is known about the action of each gene.

PatS is involved in the resolution of differentiating cell clusters into a single cell and so in our model this type of inhibition acts when a cell commits to become a heterocyst. Mitchison *et al* found a fixed inhibition range around developing heterocysts [72], so in our model committed heterocysts inhibit a range of cells on both sides from committing. We vary this range (see Figure 5.3). The PatS type inhibition ends when a cell completes differentiation and is mature.

The second type of inhibition is HetN type inhibition. HetN is described as being associated with mature heterocysts but is actually activated 6-12 hours after nitrogen deprivation [63]. In our model HetN type inhibition does not act until a period T_N after commitment. $T_N = 1$ h unless otherwise stated. We think that one hour is a reasonable minimum amount of time after commitment for the gene to begin to take effect. T_N is varied to larger values to investigate the possibility that a gap between commitment and HetN inhibition could explain the $\Delta patS$ heterocyst pattern. In our model HetN inhibition blocks the differentiation of the same range of cells as PatS inhibition. HetN inhibition never ends once it has begun.

Heterocyst maturity

In our model, committed cells always take 10 hours to mature. This time is chosen to be consistent with the range of 8-14 hours after nitrogen deprivation for commitment [26] and 18-24 hours after nitrogen deprivation for maturity and expression of nitrogenase [42]. Bradley and Carr [48] find the earliest commitment of heterocysts at 5 hours and the earliest mature heterocysts at 14 hours, which also supports a roughly 10 hour timeframe for a newly committed heterocyst to reach maturity. Mature heterocysts do not grow: their $R^{opt} = 0$. Mature heterocysts produce fixed nitrogen at a rate $G_{het} = 3.15 \times 10^6$ s^{-1} (see Section 3.2.2) after waiting a period T_S after maturity. $T_S = 1$ s unless otherwise stated. T_S is kept short in most of the situations, where we assume that it is closely tied to heterocyst maturity. We vary T_S to investigate the possibility that a gap between PatS inhibition and nitrogen fixation could be responsible for the deviation from the wild type heterocyst pattern in $\Delta hetN$ mutants, but we find little sensitivity on T_S.

Implementation

The model of Figure 5.1 is has many components and here we will explain how they are implemented. The implementations of other elements of our cyanobacterial filament model, such as the nitrogen dynamics or the cell growth, are not described here, but are similarly implemented.

There are five types of cells: vegetative cells able to commit to heterocyst differentiation, vegetative cells blocked from commitment by PatS, vegetative cells blocked from commitment by HetN, committed heterocysts, and mature heterocysts. I will describe how these cell types interact over the first 24 hours after nitrogen deprivation.

At first, all cells are vegetative cells able to commit, and these cells have nonzero cytoplasmic fixed nitrogen concentrations and grow at R^{opt}. They rapidly exhaust their cytoplasmic fixed nitrogen, but deplete their storage to continue to grow at $f_g R^{opt}$.

When storage is depleted to half its maximum amount, $0.5 f_s gl$, a cell becomes a candidate for heterocyst commitment. During every timestep, once all the candidates have been identified, one of these candidates is randomly chosen to commit to heterocyst differentiation. This committed heterocyst then inhibits its neighbours

with PatS type inhibition for a range to both sides, turning them into vegetative cells blocked from commitment by PatS. The candidates to commit to heterocyst differentiation may have changed from this inhibition, but if there are candidate cells left another is chosen randomly and the process repeated until there are no candidates left. The timestep then finishes with other tasks.

When a cell has been a committed heterocyst for a time T_N, the heterocyst also inhibits its neighbours within a range with HetN type inhibition. PatS type inhibition takes precedence and is not changed to HetN type inhibition. At the beginning of each timestep all the vegetative cells are turned into vegetative cells able to commit. Then each committed heterocyst inhibits its neighbours within a range with PatS type inhibition, turning them into vegetative cells blocked from commitment by PatS. Then committed heterocysts older than T_N inhibit their neighbours within a range, turning any vegetative cells that are not blocked by PatS into vegetative cells that are blocked from commitment by HetN. Removing and then renewing the PatS and HetN inhibition within the range each timestep allows cell division to push cells outside the inhibition range. After the PatS and HetN inhibition has been applied, the candidate cells for heterocyst commitment can be identified.

When a cell has been committed for 10 hours, the heterocyst matures and the PatS type inhibition stops, leaving only the HetN type inhibition. A time T_S after the heterocyst matures, the heterocyst begins to produce fixed nitrogen.

We explore the hypothesis that the patterning and other characteristics of heterocysts can be explained with stored fixed nitrogen, in combination with a crude model of the inhibitory action of patterning proteins PatS and HetN. This model is a significant improvement over the model of Chapter 4.

5.2.4 Some numerical details

Periodic boundary conditions were used to minimize end effects. Filaments were initiated as a single cell with zero cytoplasmic fixed nitrogen and maximum storage (i.e. $N_S = f_s g l$) in a high concentration of external fixed nitrogen. After seven days of growth in high external fixed nitrogen the external fixed nitrogen concentration was set to zero and the amount of cytoplasmic nitrogen set to 5% of the current fixed nitrogen content incorporated into the cell: $0.05 \cdot g \cdot l_i$. This amount is chosen so

that the depletion of this cytoplasmic fixed nitrogen approximately coincides with the 0.5 hour time at which nitrate assimilation genes are activated [131], one of the earliest signals of nitrogen deprivation. The time at which the external fixed nitrogen concentration is set to zero is the time of nitrogen deprivation from which all other times are measured.

Most of our measurements are straightforward. Filament length and heterocyst frequency are simply recorded every hour. Heterocyst spacing distributions are recorded at the indicated time. The fraction of the cell cycle is determined at commitment using the cell length as the indicator of cell cycle progress. Commitment is measured similar to the experimental technique of Yoon and Golden [26], where ammonia was added to the medium at a certain time to quickly revert any uncommitted proheterocysts to vegetative status and the heterocyst frequency recorded at 24 hours. In the simulation the external fixed nitrogen level was simply changed to a high concentration at a certain time, and the filament allowed to grow until the frequency was recorded at 24 hours.

5.3   Results for wild type

We first looked at the length of the filament in the first 48 hours after nitrogen deprivation, shown in Figure 5.2(a). We recover the same qualitative growth as Asai *et al* [53] (inset of Figure 5.2), with slow growth until 20-24 hours followed by increased growth. The growth we find is slower than that of Asai, and this is discussed in Section 5.4.

Heterocyst commitment is shown in Figure 5.2(b). The filament is deprived of nitrogen at zero hours and then at some later time the external fixed nitrogen is raised to a high concentration. This rise stops further commitment to differentiation and at 24 hours the cells that had committed before the rise in the fixed nitrogen level will be mature heterocysts. This is the experiment of Yoon and Golden [26] and we recover the same qualitative commitment curve as they observe (Figure 5.2(b) inset). Commitment occurs primarily from 8-14 hours after nitrogen deprivation.

The approximate commitment timing and the range of commitment is determined by the fixed nitrogen storage. Commitment does not begin until after 8 hours; this is when the first cells have run down half of their nitrogen storage. This earliest time is
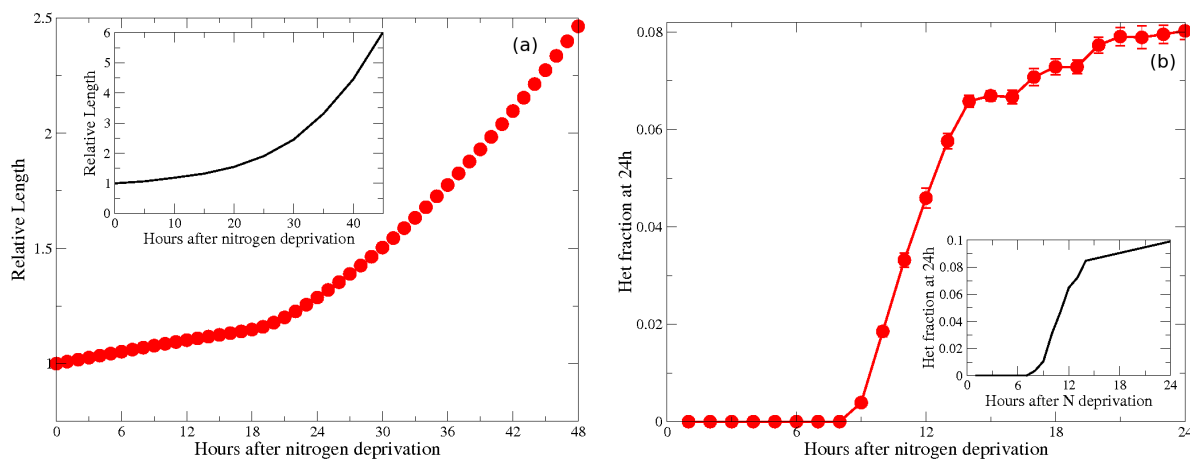
Figure 5.2: (a) Relative length of the cyanobacterial filament. The inset plot is from the length data of PCC 7120 of Asai *et al* [53] from movie B of the supplemental data. For both plot and inset the nitrogen deprivation begins at zero hours and the relative length of one is set to be the length at the onset of nitrogen deprivation. (b) The heterocyst fraction at 24 hours after initial nitrogen deprivation against time, with the time indicating when the external fixed nitrogen level was raised. The inset plot is from Yoon and Golden [26] for PCC 7120. For both plot and inset the nitrogen deprivation begins at zero hours. Model data for both (a) and (b) used a five cell inhibition range. Only qualitative agreement is expected between the experimental and modelling curves, since most parameters have been approximated.

determined by a combination of the smallest cells and the fastest growth rate. After this time, cells will continue to commit until all cells have either committed to differentiation or been blocked by PatS type inhibition. There is a range in both storage amount (due to its proportionality to cell size) and cell growth rate; these ranges combine such that commitment does not stop until about 14 hours. The heterocyst fraction increases beyond this point but that is not due to further commitment, but instead due to cell division and filament growth following later ammonia application. In the experiment, when the external fixed nitrogen level is increased, the cells that have not committed will have sufficient fixed nitrogen for optimal growth. The later the external fixed nitrogen level is increased, the fewer cells are able to divide before 24 hours to dilute the committed and mature heterocysts, and the higher the heterocyst frequency at 24 hours.

The heterocyst spacing distribution at 24 hours from the model is shown in Figure 5.3 for inhibition ranges of three to seven cells. There are two trends as the inhibition
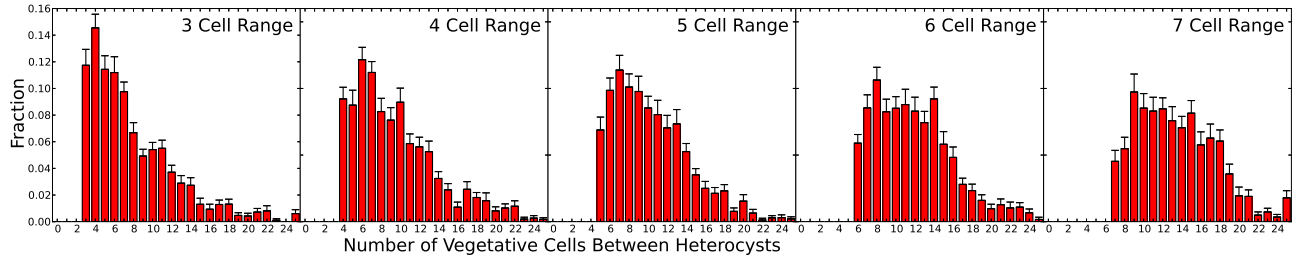
Figure 5.3: Heterocyst spacing distribution at 24 hours after nitrogen deprivation from the model, using different inhibition ranges. Inhibition ranges can overlap, leading to the nonzero counts for spacings equal in size to the inhibition range.
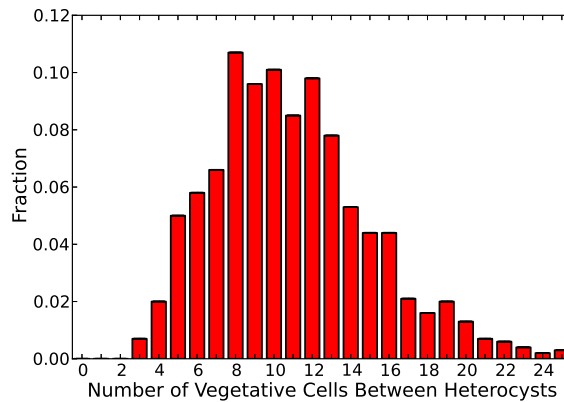
Figure 5.4: The heterocyst spacing distribution at 24 hours after nitrogen deprivation for wild type PCC 7120 from Toyoshima *et al* [46].

range increases. As expected, the minimum spacing between heterocysts increases as the inhibition range between heterocysts increases. Also, for low inhibition ranges the spacing distribution decays with the distance from the heterocyst, while at the higher inhibition ranges the distribution is more peaked near its mean. Due to the inhibition in the model having a sharp cutoff, the lower end of distribution does not increase gradually but instead does so suddenly.

The experimental heterocyst spacing distribution for PCC 7120 at 24 hours from Toyoshima *et al* [46] is shown in Figure 5.4. This distribution shows a gradual increase at the lower spacings, is fairly flat in the middle of the distribution and has a tail at larger spacings. It is also similar to the distribution for PCC 7120 at 30 hours found by Khudyakov *et al* [47].

The distribution for the five cell inhibition range is the most similar to the experimental distribution because it has the flat top and tail at the correct spacings. It does not have a gradual increase at the lower spacings because the model always has a sharp inhibition range, which does not reflect the stochasticity in inhibition range likely present during real differentiation.
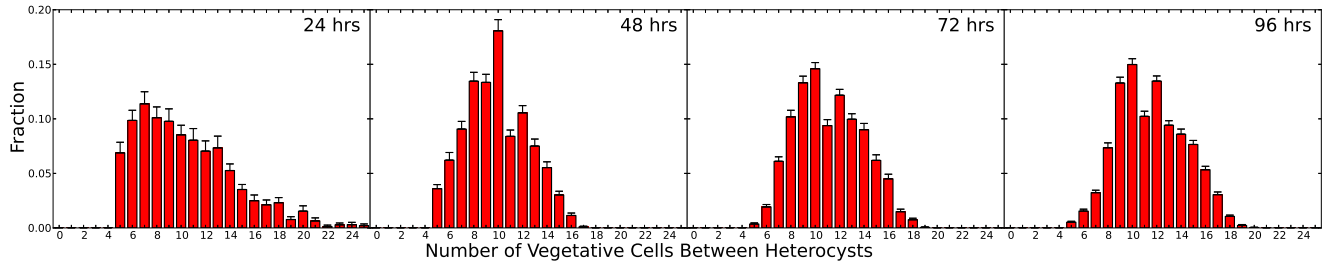
Figure 5.5: Heterocyst spacing distribution from model every 24 hours for an inhibition range of 5. Time since nitrogen deprivation in the top right corner of each histogram.
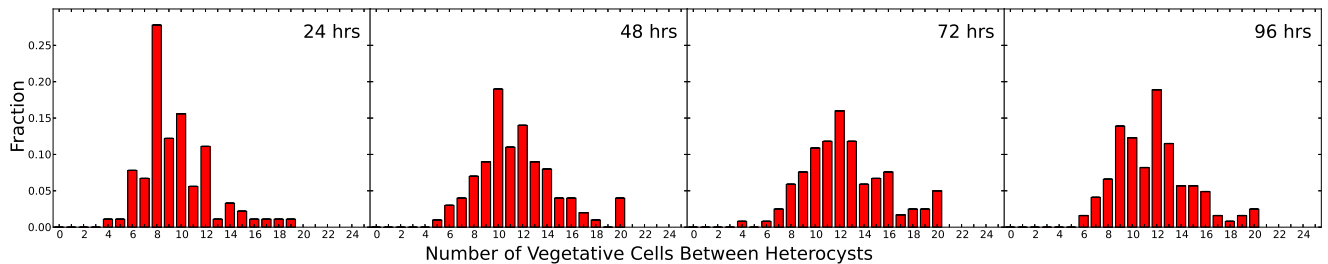
Figure 5.6: Wild type heterocyst spacing distributions for PCC 7120 from Yoon and Golden [26]. The bars at 20 represent all spacings $\geq$ 20. Time since nitrogen deprivation in the top right corner of each histogram.

The heterocyst spacing distribution from the model with an inhibition range of five for the first four days is shown in Figure 5.5. On the first day the distribution is weighted towards lower spacings, has a lower mean, and a tail at larger spacings. The distributions at later days are more symmetric, have higher means, and do not have tails for large heterocyst spacings.

Comparing to the experimental distributions from Yoon and Golden [26] in Figure 5.6, we see a similar progression to that of our model over the first four days. The first day is weighted towards lower spacings, has a low mean, and a tail for larger spacings. Later days have a more symmetric heterocyst spacing distribution with higher means.
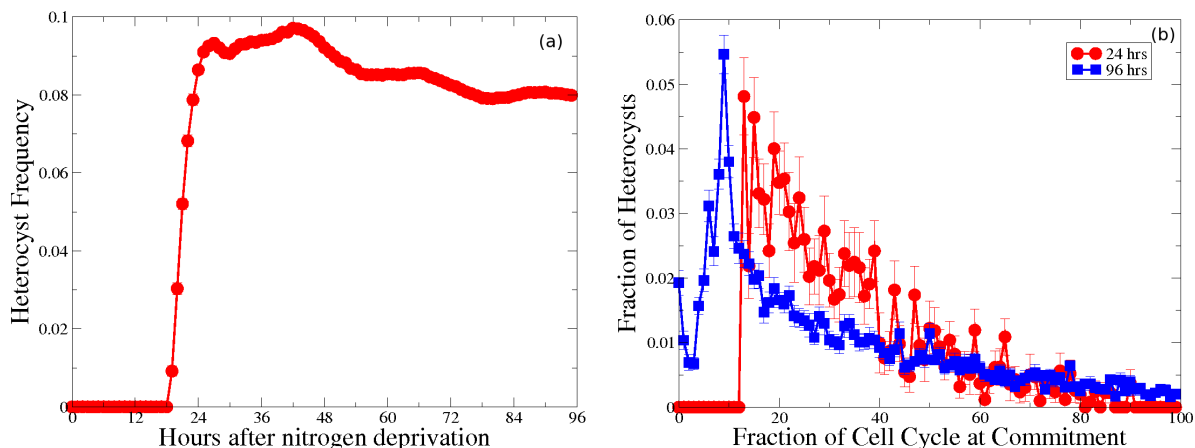
Figure 5.7: (a) is the heterocyst frequency against time since nitrogen deprivation for an inhibition range of five. (b) is the distribution of cell size of heterocysts at commitment for an inhibition range of five. The times since nitrogen deprivation are labelled in the legend.

Figure 5.7(a) shows the heterocyst frequency vs. time after nitrogen deprivation. The first heterocysts mature at 19 hours and the heterocyst frequency increases rapidly for several hours afterwards until it reaches roughly the steady state level. After about 24 hours the heterocyst frequency oscillates with approximately a 24 hour period, and frequencies dampen at later times. The frequency also drops from a peak at about 42 hours at nearly 10% to about 8% at 96 hours. The decay of the heterocyst frequency comes about because the frequency after the first generation of differentiation is determined by the inhibition range, while later it is determined by the number of heterocysts necessary to provide sufficient fixed nitrogen for the remaining vegetative cells. Since the initial differentiation provides slightly too many heterocysts, the frequency decreases at later times. The oscillation is also due to the similarity in differentiation times for the first generation heterocysts. The second generation differentiates when the gap between heterocysts is sufficiently wide; this tends to occur at roughly the same time after differentiation for each heterocyst pair. Before the second generation of differentiation, the number of cells increases, diluting the heterocysts and decreasing the frequency. When the second generation of heterocysts matures, they tend to do so at roughly the same time, pushing the heterocyst frequency up. The dilution and differentiation/maturity together result in the oscillation.

Figure 5.7(b) shows the distribution of the cell cycle at commitment for cells that have become mature heterocysts. At 24 hours the shortest cells, from about the first 10% of the cell cycle, have not become heterocysts. For cells beyond this initial 10%, there is a strong preference for shorter cells, as the likelihood of a cell committing to heterocyst differentiation to be at a given stage of the cell cycle drops for later in the cell cycle. At 96 hours this distribution has significantly changed. There is a large peak in the distribution at about 10% of the cell cycle, dropping off sharply on both sides. At 96 hours there is an even stronger preference for shorter cells to commit to heterocyst differentiation.

5.4   Discussion of wild type results

We have developed a dynamic model for cell growth, nitrogen transport, nitrogen storage, and heterocyst commitment in cyanobacterial filaments. The model is built on experimental data and has been able to reproduce many experimental observations.

As shown in Figure 5.2(a), growth after nitrogen deprivation with the model is qualitatively similar to the growth data from the experiments of Asai *et al* [53]. There is an initial, approximately one day period of slow growth when the filament is growing on stored nitrogen and heterocysts are committing and maturing. After 24 hours many heterocysts have matured and the growth rate increases. The plots indicate that the growth of the filament in the experiment of Asai *et al* [53] is greater than the growth of our model filament. We have chosen the centre of our doubling time distribution to be 20 hours, consistent with PCC 7120 [88, 95]. However, there are many instances of PCC 7120 with shorter doubling times [132, 133], specifically in the range of 6-18 hours [134]. Growth rates were affected by a change in the fraction of fixed nitrogen as storage available for growth, f_s, which we have estimated to be 0.3. A lower value will result in a decrease in the growth of the filament between heterocyst commitment and maturity, which is between roughly 10 and 20 hours after nitrogen deprivation (results not shown).

In Figure 5.2(b), commitment is measured by increasing the external fixed nitrogen level, just as it is measured experimentally [26]. The resulting commitment curve is qualitatively similar to the data of Yoon and Golden [26], with the first cells committing to heterocyst differentiation after 9 hours of nitrogen deprivation and most

of the cells committing by about 14 hours of nitrogen deprivation. The timing of commitment for a given cell is determined by the amount of fixed nitrogen storage it has. As described in Section 5.2.3, when a cell has used half of its storage, and if it is not inhibited from differentiation by a nearby committed or mature heteocyst, then it will commit to differentiation. This model predicts that strains that grow faster will commit to differentiation sooner, because they will use up the stored fixed nitrogen faster. Indeed Bradley and Carr [48] performed an experiment similar to the commitment experiment of Yoon and Golden [26] using *Anabaena cylindrica* with a generation time of 16.4 hours (shorter than the typical PCC 7120 20 hours) and found commitment ranged from 5 hours to 10 hours, instead of the 8 to 14 hours for PCC 7120 [26]. Our commitment results are affected by a change in either the storage fraction, f_s, or the fraction of storage that must be emptied before commitment, which is 0.5. Lower or higher values of either lead to earlier or later commitment, respectively.

In Figure 5.3, we compare the heterocyst spacing distributions found at 24 hours for different heterocyst inhibition ranges. We found that a five cell inhibition range was the closest match to the experimental heterocyst spacing distribution of Toyoshima *et al* [46], shown in Figure 5.4. The model was able to recover the position and tail of the distribution, but because our model used a sharp, coarse grained, inhibition range, we were unable to recover the gradual increase in the distribution for the lower heterocyst spacings. A detailed model of the genetic interactions that this coarse grained inhibition range represents would likely be able to recover all aspects of the distribution more accurately. Nevertheless, the five cell inhibition range that we found agrees with the inhibition range recorded by Mitchison *et al* [72] for *Anabaena catenula*.

Our model was also able to qualitatively recover the characteristics of the heterocyst spacing distribution at 48, 72, and 96 hours. The distributions from the model are in Figure 5.5 while the experimental distributions from Yoon and Golden [26] are in Figure 5.6. The experimental distributions at later times have a slightly higher mean, are much more symmetric, and do not have as much of a tail as the 24 hour distribution. Experimental distributions at later times from Toyoshima *et al* [46] and Khudyakov *et al* [47] also follow this trend. The modelled results recover this change.

The heterocyst spacing distribution changes from an asymmetric to a symmetric

distribution from 24 to 48 hours. This is because the 24 hour spacings are controlled by the lateral inhibition range and this causes smaller spacings to be favoured. The 48 hours spacings are instead controlled by the fixed nitrogen production of the heterocysts. This production prevents cells near to the heterocysts from starving and committing to heterocyst differentiation. The maximum in the heterocyst spacing distribution is defined by the maximum number of cells away a cell can be before it will differentiate into a heterocyst, and the minimum is half of the minimum number of cells away a cell can be before it will differentiate into a heterocyst. This leads to the spacings to be more evenly favoured and a more symmetric distribution.

Figure 5.7(a) shows the heterocyst frequency in time. The frequency is zero until approximately 20 hours, when it rapidly increases as the first generation of heterocysts matures, reaches a maximum, and then slowly decreases and oscillates. The mature heterocyst frequency has been observed to reach a maximum after the first heterocyst generation matures and then decrease [12]. In experiments the heterocyst frequency has also been observed to oscillate [48].

The literature holds several suggestions that the cell cycle is important to the differentiation of heterocysts [31, 73, 129], and at least one that points to a requirement that cells be early in the cell cycle to be able to commit to heterocyst differentiation [72]. The existence of these suggestions in the literature have been sufficient to warrant the explicit demonstration of heterocyst differentiation without recent cell division [46, 53]. Figure 5.7(b) shows a preference for the commitment of shorter cells, demonstrating that the cell cycle affects heterocyst commitment. This is despite our model including no cell cycle criteria in determining heterocyst commitment. The cell cycle effect that we find is instead due to the role of storage in the heterocyst commitment model. In our model the fixed nitrogen storage is proportional to the size of the cell. A small cell growing at the same rate as a larger cell will run out of storage before the larger cell. Smaller cells therefore commit to becoming heterocysts more often than larger cells.

5.5   Results for genetic knockouts

Figure 5.8 shows the heterocyst spacing distributions for filaments with the PatS type inhibition removed from the model. The amount of time after commitment that
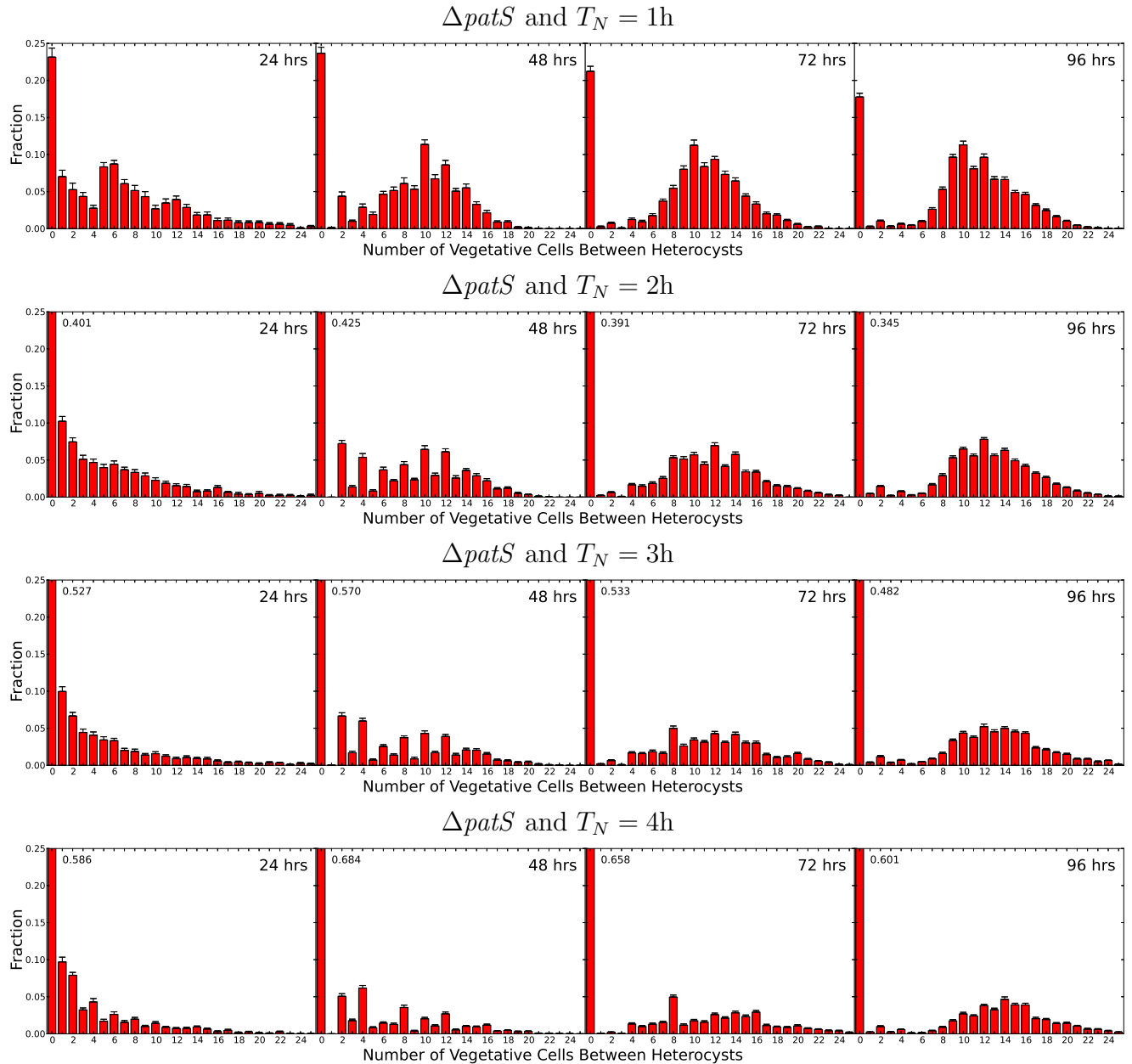
Figure 5.8: The heterocyst spacing distribution from the model without PatS inhibition for proheterocysts but with HetN inhibition. T_N is the amount of time after commitment that HetN type inhibition begins (see Figure 5.1). Each row shows the patterns at 24, 48, 72, and 96 hours for a different T_N, as indicated.

HetN type inhibition begins, T_N, is varied from one to four hours. There are many striking differences in common for all four T_N, with PatS inhibition removed, when compared with the wild type model of Figure 5.5. With PatS inhibition removed, spacings of zero, known as multiple contiguous heterocysts (Mch), occur and are the

most common spacing for all four days and all four T_N values. The number of Mch is fewest for $T_N = 1$ hour and increases as T_N is increased. In contrast, the wild type model does not have Mch. There are also nonzero but low heterocyst spacings present without PatS inhibition that are not present for the wild type. For the $T_N = 1h$ the 24 hour spacing distribution shows some structure, while the longer waiting times all decay from a spacing of zero. With all four T_N values the distribution is furthest from the wild type for 24 hours and recovers somewhat toward the wild type distribution in subsequent days. This recovery is strongest for $T_N = 1h$ and weaker as T_N increases, likely due to an increasing proportion of heterocysts as Mch for longer T_N. The number of Mch follows a trend across all four T_N of increasing from the first to second day, and then decreasing in subsequent days.
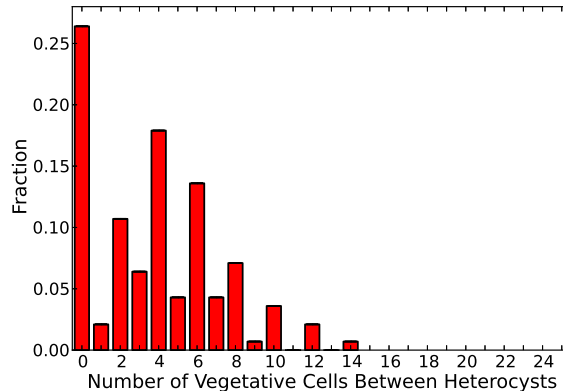
Figure 5.9: The heterocyst spacing distribution at 24 hours after nitrogen deprivation for a $\Delta patS$ strain from Yoon and Golden, 1998 [6]. The data is continued to later times in Figure 5.10, although without Mch data.

Our modelled distributions lacking PatS inhibition can be compared to experimental $\Delta patS$ mutants lacking a functional *patS* gene. Figure 5.9 shows such data from Yoon and Golden in 1998 [6] and Figure 5.10 from Yoon and Golden in 2001 [26]. The 1998 data is only for 24 hours. The 1998 data indicates a little more than 25% Mch spacing, similar to the almost 25% Mch spacing for $T_N = 1h$ with the model when PatS inhibition is removed in Figure 5.8(a). The 1998 data also shows that the distribution for nonzero but low spacings has structure, again supporting the one hour HetN wait time after commitment. The 2001 data shows that although the 24 hour data is far from the wild type, that the distribution somewhat recovers a wild type

distribution for later days, as do the distributions from the model. Unfortunately there is no experimental data on the number of Mch spacings as time passes.
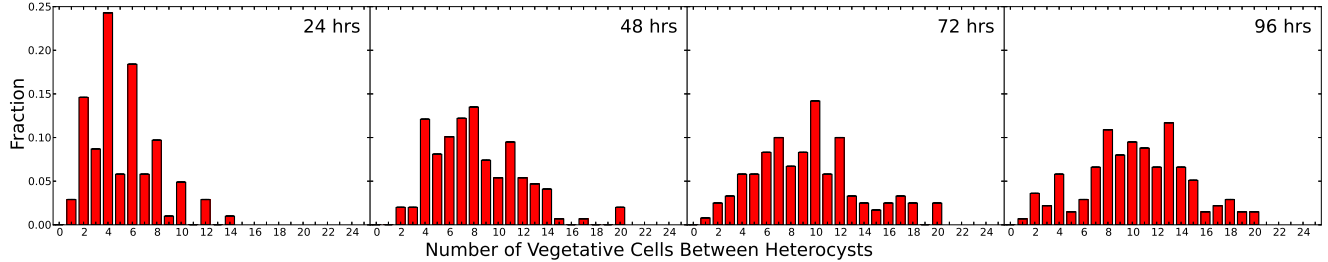
Figure 5.10: The heterocyst spacing distribution for a $\Delta patS$ strain from Yoon and Golden, 2001 [26]. Time since nitrogen deprivation in the top right corner of each histogram. The paper did not include quantitative data for Mch (number of vegetative cells between heterocysts of zero).

Figure 5.11 shows the heterocyst spacing distributions for filaments with the HetN type inhibition removed from the model. T_S is varied from one minute to four hours. At 24 hours after nitrogen deprivation, there is little difference between the distributions lacking HetN inhibition and the wild type distributions of Figure 5.5, although the biggest differences are with the longer T_S. The main difference between wild type and filaments lacking HetN inhibition is at later times. For 48 hours and afterwards the filaments without HetN inhibition have Mch and short spacings between heterocysts. The value of T_S, in the range of one minute to four hours, has little effect on these later time distributions.

Here we draw a distinction between two possible types of multiple contiguous heterocysts. In the systems lacking PatS, the Mch occur when neighbouring cells commit to heterocyst differentiation before either of them mature. They cannot commit beside a mature heterocyst because HetN prevents commitment beside a mature heterocyst. The Mch for systems lacking HetN occur by having a cell commit to heterocyst differentiation beside a pre-existing heterocyst that has not yet started to produce fixed nitrogen (see the $\Delta hetN$ gap in Figure 5.1), as PatS prevents the commitment of two neighbouring cells.

Although there are no experimental heterocyst spacing distributions for $\Delta hetN$ filaments there are experimental observations with such systems. $hetN$ mutants exhibit a wild type heterocyst pattern after 24 hours of nitrogen deprivation, but have an Mch phenotype after 48 hours [61]. This is what we see in the results using the

$\Delta hetN$ and *de novo* fN production waits 1min after PatS is off

$\Delta hetN$ and *de novo* fN production waits 10min after PatS is off

$\Delta hetN$ and *de novo* fN production waits 1h after PatS is off

$\Delta hetN$ and *de novo* fN production waits 4h after PatS is off

Figure 5.11: The heterocyst spacing distribution from the model without HetN inhibition but with PatS inhibition.

model.

5.6 Discussion of gene knockout results

Yoon and Golden have found that knocking out PatS results in heterocyst spacing distributions that are not wild type [6, 26], and this has been attributed to PatS
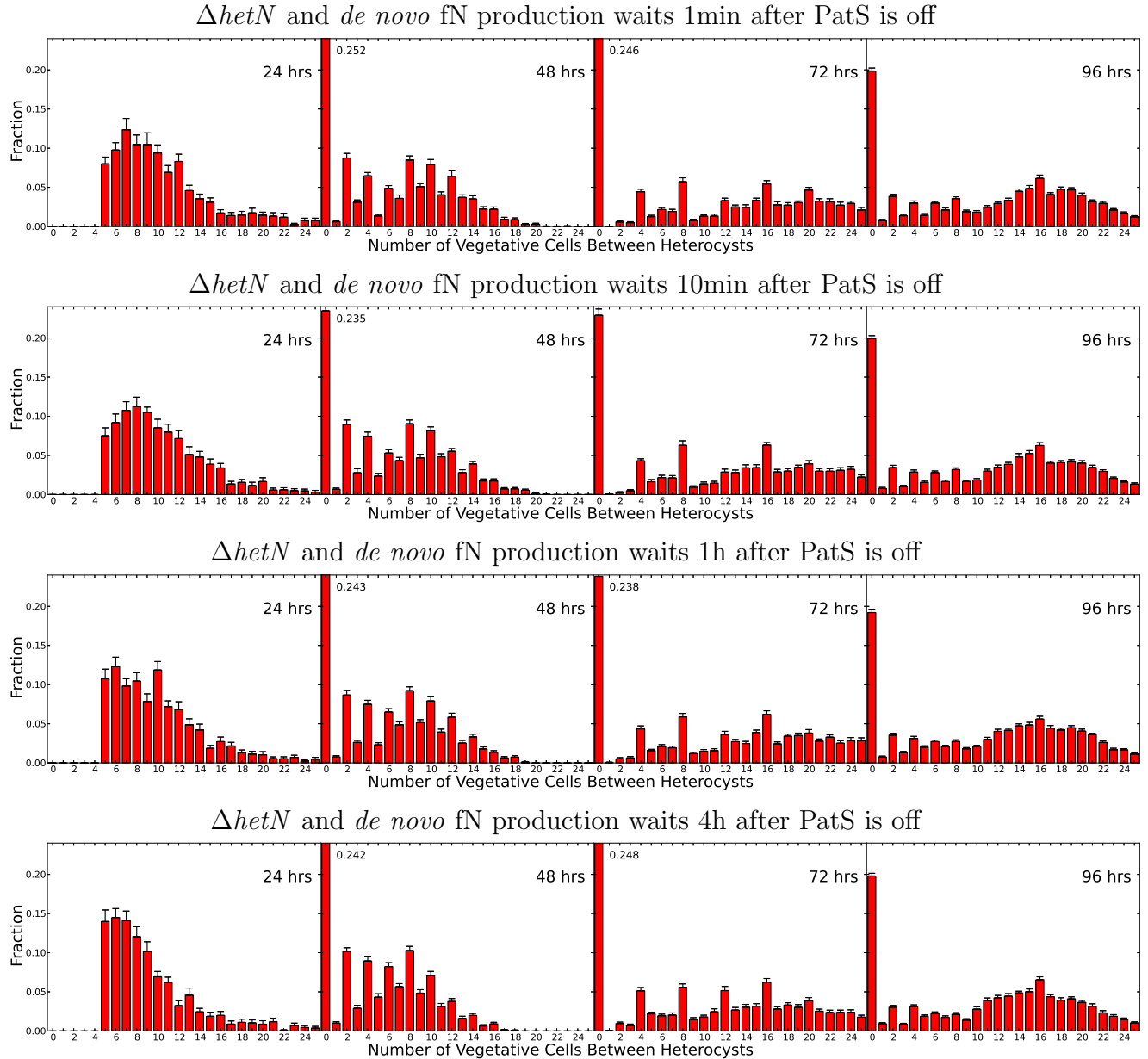
being a diffusible inhibitor of heterocyst differentiation. They also found [26] that the pattern somewhat recovers at later times, even without functional PatS. By removing the PatS type inhibition of heterocyst differentiation from our model, we were able to produce qualitatively similar distributions, as shown in Figure 5.8. The time that *hetN* is first expressed has only been determined to be in the range of 6-12 hours after nitrogen deprivation [63]. Assuming that *hetN* is not expressed until after a cell has committed to heterocyst differentiation, we varied the amount of time after commitment that HetN type inhibition would begin to take effect, in systems with the PatS inhibition removed, and found that a one hour waiting time was most consistent with the experimental distributions.

Knocking out *hetN* results in wild type heterocyst spacings at 24 hours after nitrogen fixation, but significantly different spacing distributions at later times [61] with Mch and short vegetative cell intervals between heterocysts. We are able to recover this phenotype using our model by removing the HetN type inhibition and having a temporal gap between the end of PatS type inhibition and the beginning of nitrogen fixation, as shown in Figure 5.11. We vary the amount of time between the end of PatS inhibition and nitrogen fixation to as little as one minute and still observe Mch and short heterocyst spacings. In the experiment of Wei *et al* using PCC 7120 [135], the first mature heterocysts were observed between 15 and 16 hours after nitrogen deprivation, but the nitrogenase genes *nifHDK* were not detected until 18 hours, supporting the idea of a gap between maturity and nitrogen fixation.

5.7   Predictions

We have shown that the timing of heterocyst commitment can be explained by the depletion of stored fixed nitrogen. The depletion would occur earlier in instances where the same amount of storage is initially present but the growth rate was higher. We therefore predict that if the growth rate of PCC 7120 could be controlled (say, by variation of light intensity or phosphate concentration), then the timing of commitment to heterocyst differentiation could be changed, with faster growing strains committing earlier. Also, if cyanophycin was genetically knocked out (such as in the strain studied by Ziegler *et al* [66]), we expect that the commitment to heterocyst differentiation would occur earlier and more uniformly. Specifically we expect that

hetR and *patS* expression, possibly reported by a GFP fusion, would occur earlier (do not have to wait for storage to run out) and more uniformly (cells do not have different amounts of stored fixed nitrogen) in strains lacking cyanophycin compared to the wild type with cyanophycin.

The cell cycle effect we observed using our model depended upon shorter cells having less stored fixed nitrogen, leading to these shorter cells being more likely to deplete their storage first and commit to heterocyst differentiation. Genetically knocking out cyanophycin would then lead to a weaker or nonexistent cell cycle effect in comparison to wild type strains. We also predict, consistent with Figure 5.7(b), that the cell cycle effect will be stronger in later generations of heterocysts than in the first generation of heterocysts.

We were able to reproduce the Mch phenotype in $\Delta hetN$ mutants for 48 hours and later by having a temporal gap between the end of PatS inhibition and the beginning of nitrogen fixation. In the wild type, HetN inhibition would prevent heterocysts from committing to heterocyst differentiation during this time, but without HetN the starving cells are free to commit. We predict that the Mch in filaments lacking functional *hetN* will occur by having cells commit to heterocyst differentiation beside mature heterocysts, rather than two neighbouring cells committing and maturing around the same time, as this second option would be prevented by PatS inhibition.

Chapter 6

Percolation of Starvation

6.1 Motivation and background

The previous three chapters have focused on the cyanobacterial filament in the context of heterocysts. In this chapter I want to examine the patterns of starving cells within the filament in the absence of heterocysts. Cyanobacterial filaments without heterocysts are real systems. Many strains of filamentous cyanobacteria do not have heterocysts [22], and heterocystous strains with $\Delta hetR$ cannot differentiate heterocysts.

Starving cells in one dimension, or in two dimensions with time as the second dimension, can be considered with percolation theory. Percolation theory can be used to consider the pattern of occupation of squares in a grid, and if we assign starving cells as occupied and non-starving cells as unoccupied, then we can explore the patterns of starving cells using the tools of percolation theory.

We want to investigate the transition of the filament from starving to non-starving as the external fixed nitrogen level increases. Such transitions in percolation theory involve finding critical exponents that characterize the behaviour and patterns of the occupied squares near the transition. Percolation usually involves the assumption that the occupation of a square is uncorrelated with the occupation of other squares. This assumption prevents an exact mapping from our model to percolation and we will investigate to what extent percolation theory does apply and how the results from the model deviate from percolation theory.

6.2 Percolation background

The information for this background is largely from *Introduction to Percolation Theory* by Stauffer and Aharony [136].

Percolation can be considered in any number of dimensions and with any lattice

geometry. However, in this background we will only consider square lattices in one and two dimensions. Beginning with a two dimensional square lattice, each square in the lattice is either filled in or not filled in. Squares that are filled in are referred to as occupied and those that are not are referred to as not occupied. A cluster is defined as a group of neighbouring occupied squares. To be considered neighbours the squares need to be nearest neighbours, i.e. they must have one side in common. Squares that touch only at a corner are not considered nearest neighbours. The squares within a cluster are connected to one another by an unbroken chain of nearest-neighbour links.

In percolation theory the occupied sites are usually not considered to be correlated. The occupation of each square is random, with p being the probability of a square being occupied. A cluster is said to percolate in a system if the cluster extends from the top to the bottom and the left to the right of the system. In the thermodynamic limit such a cluster would be infinite and a percolating cluster can be referred to as an infinite cluster. As the occupation probability p is increased from zero towards one, at some point it will reach the critical concentration p_c, which is the lowest p value at which a percolating cluster is formed in an infinite system.

The measurement we are going to consider is the number of clusters of size s, n_s. It is found that the function $v_s(p) = n_s(p)/n_s(p_c)$ depends only on the variables s and $(p - p_c)$ and only in the combination $|p - p_c|s^\sigma$. Power law behaviour is also expected at the critical point. We then assume that the distribution is

$$n_s = s^{-\tau} f[(p - p_c)s^\sigma]. \tag{6.1}$$

The Fisher exponent τ describes the power law term and is a constant. f is a scaling function that can act as a cutoff for the power law. σ is a free exponent, and if $|(p - p_c)s^\sigma| \ll 1$ then $f[(p - p_c)s^\sigma]$ approaches a constant value, leaving a power law. Thus, close to p_c we generally expect a power law distribution of cluster sizes.

6.2.1 One Dimensional Lattice

On a one dimensional lattice all sites must be occupied for a cluster to reach both ends of the lattice. This means that the critical concentration $p_c = 1$.

The probability of a cluster of s occupied sites beginning on a particular site is equal to the product of p^s, the probability of a single site being occupied to the power

of the number of sites, and $(1-p)^2$, the probability of a single site being unoccupied squared for the unoccupied sites on each end. This gives us

$$n_s = p^s(1-p)^2 \tag{6.2}$$

We want to find a power law in s for n_s near p_c and so we wish to find n_s as a function of s alone and independent of p for $p \to p_c$.

We can rewrite $p = e^{ln(p)} = e^{ln(p-1+1)}$. $ln(1+x) \approx x$ if x is small, so if p is near to one then $p \cong e^{p-1}$. $p_c = 1$ (making the condition of p near one into $p \to p_c$) and so $p = e^{p-p_c}$. Close to $p = p_c$, we can rewrite Equation 6.2 as

$$n_s(p) = (p_c - p)^2 e^{-(p_c-p)s} \tag{6.3}$$

Considering Equation 6.1, we know the scaling function has as its argument $(p-p_c)s^\sigma$. Defining $z \equiv (p-p_c)s^\sigma$, we can rewrite the right side of Equation 6.3 as $s^{-2\sigma}z^2 e^{zs^{1-\sigma}}$. The scaling function f must be a function of z alone, so $e^{zs^{\sigma-1}}$ demands that $\sigma = 1$. This gives us the scaling form equation

$$n_s(z) = s^{-2}z^2 e^z, \tag{6.4}$$

which satisfies Equation 6.1 with $\tau = 2$. Knowing that $f(z) = z^2 e^z$ goes to a constant if $|(p - p_c)s^\sigma| \ll 1$, when $p \to p_c$, n_s has a power law in s for one dimensional percolation when near p_c. Also, $z = (p - p_c)s$ is negative because $p \le p_c$ and is causing a decay with the increasing s.

6.2.2 Two Dimensional Lattice

For a two dimensional square lattice $p_c = 0.5927$, found by Monte Carlo simulation [137]. For the power law in n_s, it has been found exactly that $\tau = 187/91 \approx 2.055$ for two dimensions.

6.3 Model

This chapter has used the model from Chapter 4, described in Section 4.2, with 1% leakage and no heterocysts.

Many of our measurements are made of starving cells. Starving cells are defined as cells that have a cytoplasmic nitrogen level of zero, meaning the cell is growing slower

than its optimal growth rate. The measurement of the fraction of cells starving is simply the number of starving cells divided by the total number of cells in the filament. For one dimension, cluster length is measured. The length of a starving cluster is defined to be the number of consecutive cells in the filament that are starving.

For two dimensions (space and time), cluster size is measured. Consecutive starving cells in the filament at a specific time are part of a cluster. A starving cell and that same cell starving at the next time sampled are also part of a cluster. If a starving cell divides and one or both of the daughter cells are starving, then those two or three cells are part of a cluster. Cells that can be connected to one another through these spatial or temporal links are a single space-time cluster.

For two dimensions, the average cluster size is defined as the $\sum_{s=1}^{\infty} s \cdot n_s / \sum_{s=1}^{\infty} n_s$, where n_s is the number of clusters of size s.

6.4   Percolation
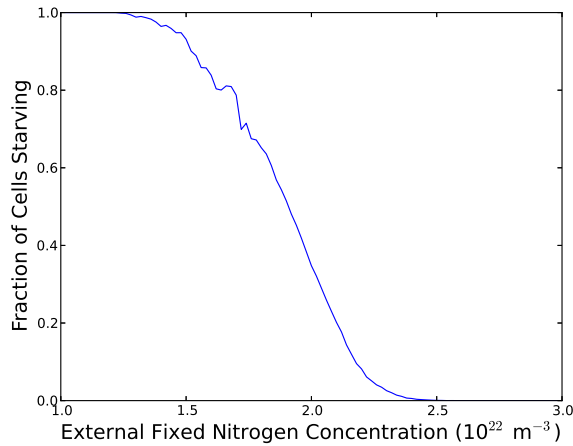
Figure 6.1: Fraction of starving cells as the external fixed nitrogen concentration varies. Fraction first decreases below one at 1.22×10^{22} m^{-3} and reaches zero at 2.66×10^{22} m^{-3}. The system used the model described in Section 4.2, with $D_C = 1.54$ μm s^{-1}, $D_O = 2.9 \times 10^{-18}m^3/(\mu$m·s$)$, $D_M = 0.01 D_O/(\pi \ \mu$m$^2)$, $T_D = 20$ h, $\Delta = 4.5$ h, $l_{min} = 2.25$ μm, and $g = 6.2 \times 10^9$ μm^{-1}.

Only when the fraction of cells starving is between zero and one can the starving cells form clusters in the system. Figure 6.1 shows the fraction of cells starving in the filament as a function of the external fixed nitrogen concentration.

As discussed in Section 6.1, we are using a model with no heterocysts, and so the only source of fixed nitrogen for growth is from the import of external fixed nitrogen. The import of fixed nitrogen is linearly related to the length of the cell (see Section 4.2.1). For low concentrations of external fixed nitrogen, no cell is long enough to import enough fixed nitrogen to fully supply growth and not starve. For high concentrations of external fixed nitrogen, all cells are long enough to import enough fixed nitrogen to fully supply growth. For intermediate concentrations of external fixed nitrogen, some of the cells are long enough to import enough fixed nitrogen to fully supply growth. Whether or not a particular cell is able to obtain enough fixed nitrogen is dependent on how long the cell is and its growth rate, because faster growing cells require more fixed nitrogen to avoid starvation. Also, if a neighbouring cell is not starving and has extra fixed nitrogen, diffusion could lift a cell from starvation. This is discussed quantitatively in Section 4.3.

6.4.1   One dimensional percolation

We consider one dimensional percolation of starving clusters at a given time.

Figure 6.2(a) shows the cluster length distribution at various external fixed nitrogen concentrations in log-log format, which will highlight any power law behaviour. The range shown in Figure 6.2(a) covers most of the concentrations that lead to some starving cells. The best power law is seen for the second highest curve in the figure, which has $\tau = 2.048$ and starvation of 94.8% of the cells. This τ value is very close to the expected value of 2 for one-dimensional percolation, and a power law with $\tau = 2$ agrees well in Figure 6.2(a) (dashed line). The highest curve in the figure has an even higher percentage of starving cells, and we would expect that this larger percentage, being closer to the critical concentration of 100%, would yield an improved power law. This is not the case and could be caused by correlations of the starving cells or could be an artefact of the finite system size.

Power law behaviour is seen for the highest curves in Figure 6.2, for which the fixed nitrogen concentration causes nearly all cells to starve. As the external fixed nitrogen concentration increases and the fraction of cells starving decreases, the curves (the lower curves in Figure 6.2(a)) begin as power laws at low cluster size but trail down for higher cluster sizes. This is expected as the exponential in Equation 6.4 becomes
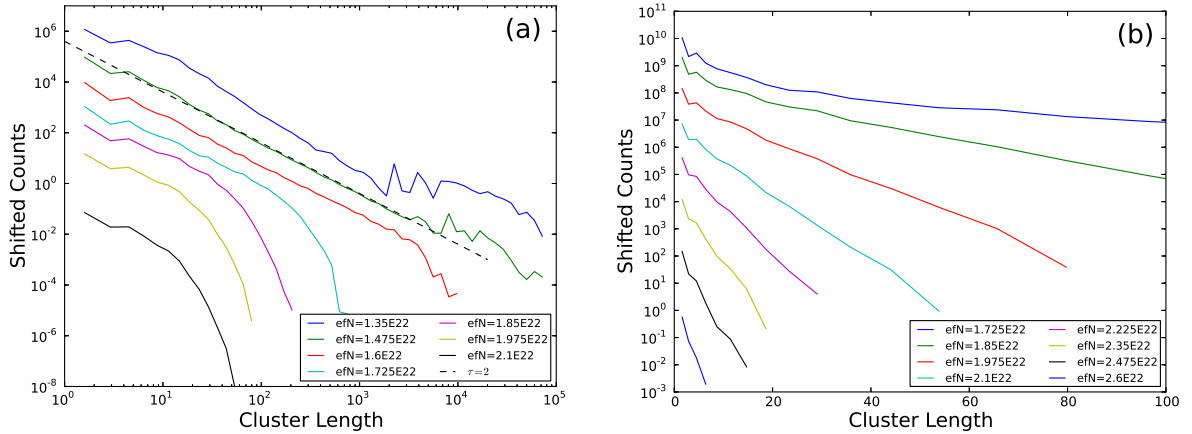
Figure 6.2: (a) Cluster length distribution in log-log format, using the same parameters as Figure 6.1. Each line is from a different external fixed nitrogen concentration, efN. The highest line is the lowest efN. The slope of the straight part of the green line for 1.475×10^{22} m^{-3} is -2.048, found by linear regression. This concentration has 94.8% starving cells. The black dashed line is a power law $s^{-\tau}$ with $\tau = 2$. The highest line has counts multiplied by 10, and each subsequent line has counts multiplied by a factor of 10 less to keep the lines from being on top of one another; this is what is meant by 'shifted' counts. (b) Cluster size distribution in log-linear format. The highest line is the lowest efN. The highest line has counts multiplied by 10000, and each subsequent line has counts multiplied by a factor of 10 less.

dominant at larger cluster sizes and starving cell densities are further from one.

Figure 6.2(b) shows the cluster length distribution at various external fixed nitrogen concentrations in log-linear format, which will highlight any exponential behaviour. The fixed nitrogen concentrations in Figure 6.2(b) are not close to $p = p_c = 1$ and so exponential behaviour is expected and observed. The slope is increasing in magnitude as the external fixed nitrogen concentration increases, as is expected because Equation 6.4 goes as $e^{(p-p_c)s}$ and p will be further from $p_c = 1$ for higher external fixed nitrogen concentrations.

6.4.2   Two dimensions

We now consider two dimensional percolation of starving clusters. The first dimension is filament length and the second is starvation time. A single cluster is any connected spacetime area of starving cells.
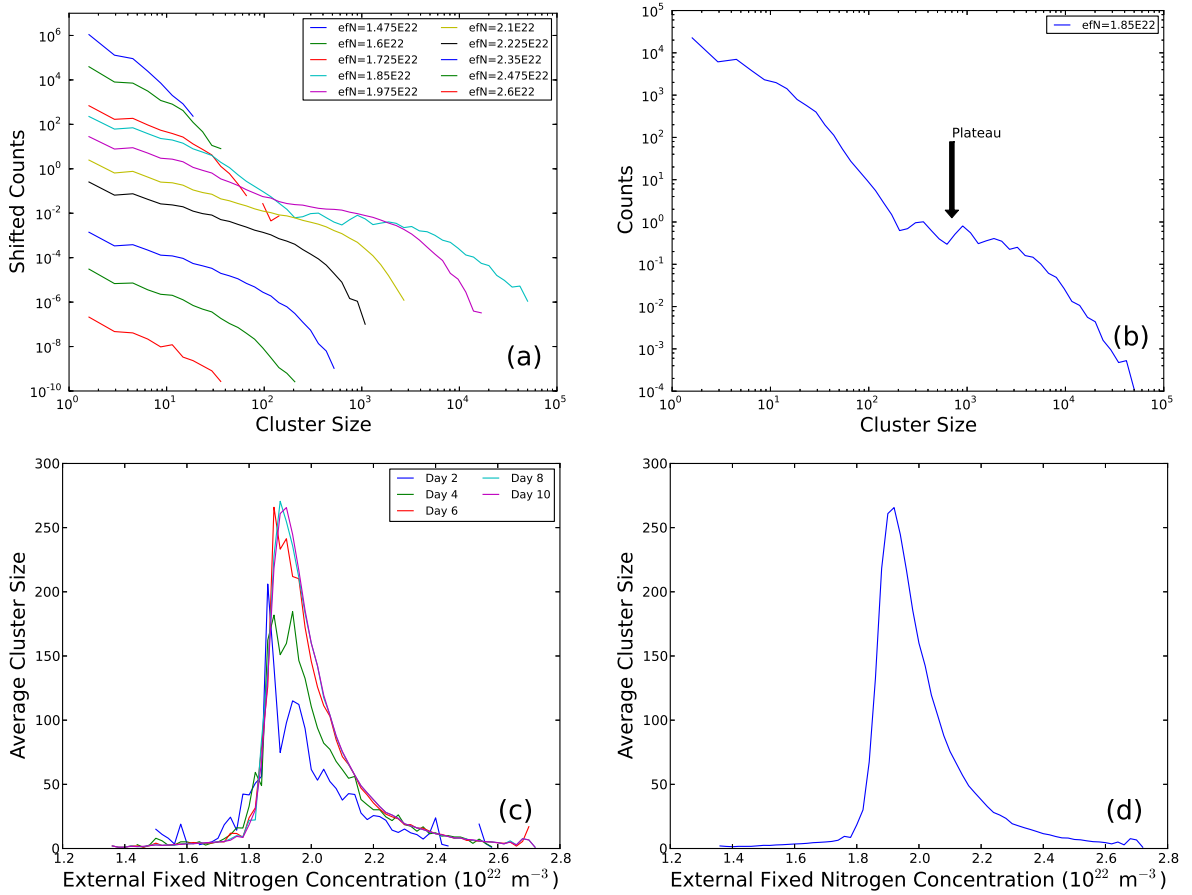
Figure 6.3: (a) Cluster size (two dimensions) distribution in log-log format, using the same parameters as Figure 6.1. The highest line has counts multiplied by 1000, and each subsequent line has counts multiplied by a factor of 10 less to keep the lines from being on top of one another; this is what is meant by 'shifted' counts. (b) The distribution with the largest cluster sizes from (a), $\rho_{efN} = 1.85 \times 10^{22}$ m^{-3}. The plateau in the distribution is indicated. (c) Average cluster size as the external fixed nitrogen concentration varies. The average does not include the 'infinite' cluster that spans the system in systems for which it exists. The day number indicates how long after the beginning of the simulation the data was taken. (d) Same as (c) except only the final distribution.

Figure 6.3(a) shows cluster size distribution in log-log format for a range of external fixed nitrogen concentrations. The lower curves are for filaments that only have a few starving cells. At those external fixed nitrogen concentrations there are not enough starving cells to form one infinite cluster and instead there are smaller, isolated clusters. These curves have flat power law regions for smaller cluster sizes and then tail off at larger cluster sizes. As more cells starve, the tailing off behaviour

is pushed out to larger cluster sizes. A spacetime visualization of this type of system is shown in Figure 6.4(a).
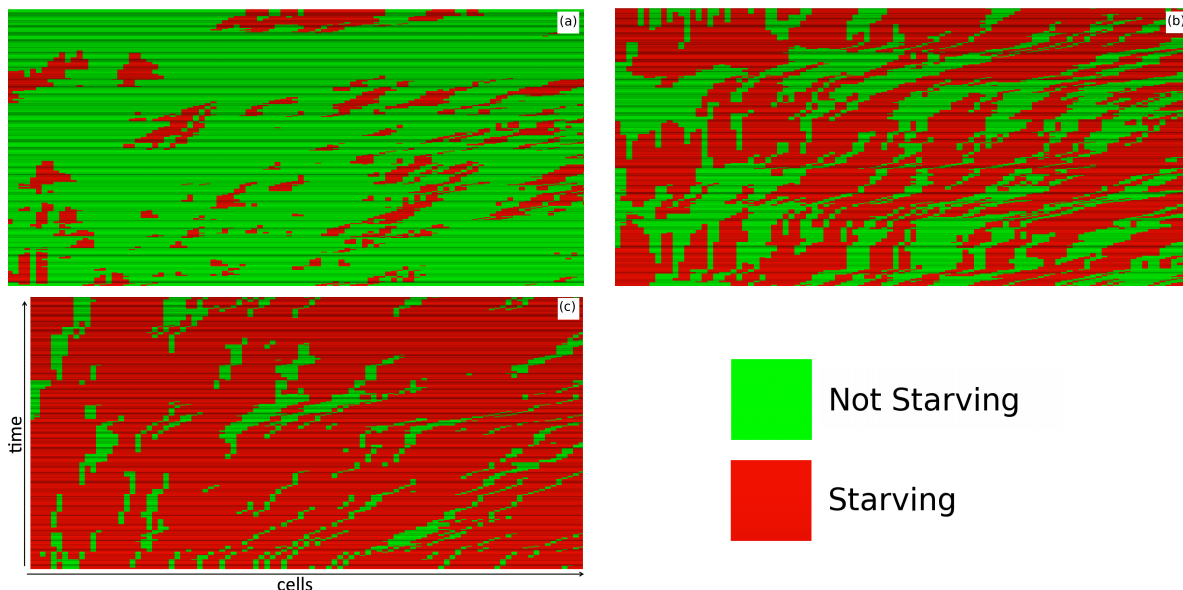
Figure 6.4: Spacetime visualizations of filament starvation, using the same parameters as Figure 6.1. Red represents starving cells and green represents non-starving cells. Each horizontal bar is the first 100 cells in the filament, and subsequent bars are separated by 10 minutes. There tends to be streaming to the right due to growth of the filament. (a) is for $\rho_{efN} = 1.6 \times 10^{22}$ m^{-3}, well below the transition of starving cells to an infinite cluster, with only a few starving cells. (b) is for $\rho_{efN} = 1.85 \times 10^{22}$ m^{-3}, roughly at the transition of starving cells to an infinite cluster. (c) is for $\rho_{efN} = 2.1 \times 10^{22}$ m^{-3}, well above the transition of starving cells to an infinite cluster, with many starving cells.

As the external fixed nitrogen concentration is decreased (higher curves in the plot) the number of starving cells increases and there are more clusters and the clusters tend to be larger. When enough cells are starving a single infinite cluster will form spanning the entire space-time system. A spacetime sample of this type of system is shown in Figure 6.4(b). This transition can be seen in Figure 6.3(d) as the peak in the average cluster size distribution (the peak is not divergent because of finite size effects, shown by the sharpening of the curve in Figure 6.3(c) as time passes and the system gets larger). Unlike the one dimensional case above, we do not observe power law behaviour at all cluster sizes when approaching this transition. For the ρ_{efN} concentration that yields the largest clusters, we instead see a plateau form in

the cluster size distribution, indicated in Figure 6.3(b). This will be discussed more below.

When the infinite cluster forms all of the cells are not starving and inside the infinite cluster there are pockets of cells that are not starving. Within these pockets of non-starving cells there can be clusters of starving cells which are not connected to the infinite cluster. A spacetime sample of this type of system is shown in Figure 6.4(c). These clusters compose the distribution for the highest curves in Figure 6.3(a).

6.4.3   Deviations from percolation in two dimensions

The lack of a power law and the plateau in two dimensions at the transition indicates that the system has correlations, which must be either spatial or temporal. In Section 6.4.1 we examined the behaviour along the filament and there are no apparent spatial correlations that could explain the plateau in Figures 6.3(a) and (b).
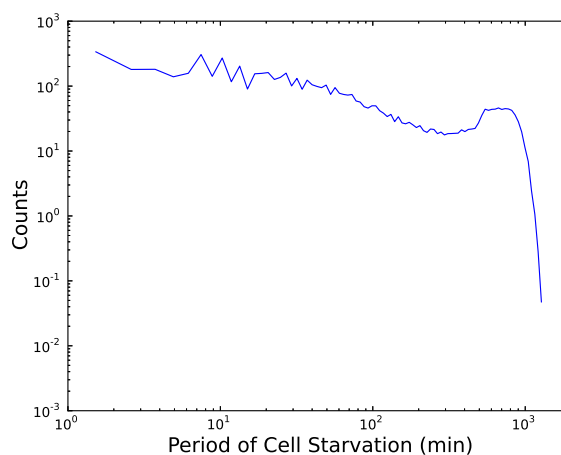
Figure 6.5: The distribution of the amount of time each cell starves for $\rho_{efN} = 1.85 \times 10^{22}$ m^{-3}, the concentration from the curve from Figure 6.3(b), using the same parameters as Figure 6.1.

Figure 6.5 shows the distribution of the amount of time each cell starves. The distribution is approximately a power law for the smaller cluster sizes, a bump below 10^3 minutes, and a cutoff above 10^3 minutes. The bump corresponds to the average cell period of 20 hours or 1200 minutes and indicates that there are temporal correlations in cell starvation associated with the cell period. From this, we expect that

the cause of the plateau in Figures 6.3(a) and (b) is due to temporal correlations in starvation associated with the cell cycle.
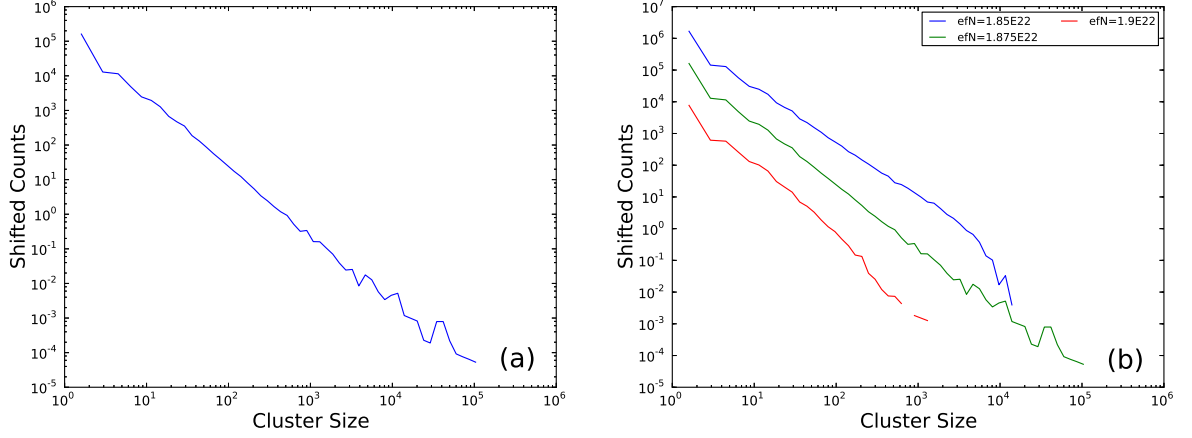
Figure 6.6: Cluster size distribution in log-log format for a bare model reduced to percolation, using the same parameters as Figure 6.1. (a) The external fixed nitrogen concentration is $\rho_{efN} = 1.875 \times 10^{22}$ m^{-3}. The slope of the line is -2.064, found by linear regression. (b) is the curve from (a) as well as two curves with nearby external fixed nitrogen concentrations. The curve with lower efN decays at high cluster size and the curve with higher efN does not record cluster sizes much higher than 10^3.

To further investigate why there is not power law behaviour at the transition to the infinite cluster we begin by stripping away several aspects of the model to leave only a two dimensional percolating system. All growth rates are set to zero and there is no cell division. All cell sizes are the same and the fixed nitrogen consumption rate of each cell is chosen every 10 seconds. There are very high losses and no exchange between cells. Together, these changes provide a bare model with a constant number of cells, each of which is starving independently of the starvation of itself and other cells in space and time. Each cell will be starving with some independent probability, corresponding to two dimensional percolation. The result at $\rho_{efN} = 1.875 \times 10^{22}$ m^{-3} is shown in Figure 6.6(a), where we have a power law with a slope of -2.064, very close to the expected power law from Section 6.2.2 of slope -2.055. Cluster size distributions for higher and lower ρ_{efN} exhibit cutoff behaviour, as seen in Figure 6.6(b), indicating that $\rho_{efN} = 1.875 \times 10^{22}$ m^{-3} is the critical value. $\rho_{efN} = 1.875 \times 10^{22}$ m^{-3} also corresponds to $p = 0.62 \pm 0.02$ for a 1000x1000 system (1000 cells, 1000 time samples), which is within two standard deviations of the expected result for criticality of $p_c =$

0.5927.

Now we want to individually add back the aspects of the model that have been removed and see the effect of each on the cluster size distribution. For each of the distributions I have adjusted ρ_{efN} to get the cluster size distribution as close to a power law as possible, judging by inspection.
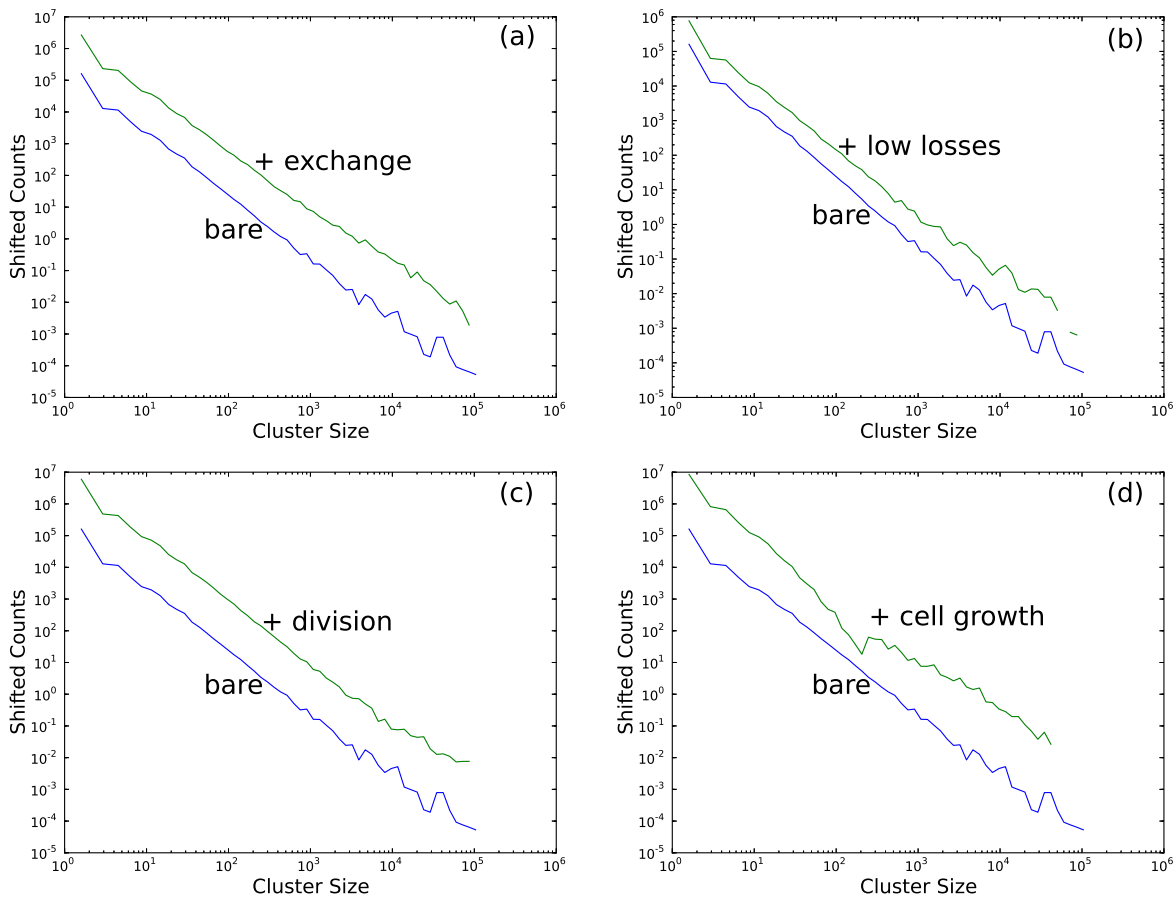
Figure 6.7: Cluster size distribution in log-log format, using the same parameters as Figure 6.1. All lower curves are for the bare model reduced to percolation with $\rho_{efN} = 1.875 \times 10^{22}$ m^{-3}. Upper curves have different aspects of the model reinserted. All curves have ρ_{efN} varied to get as close as possible to power law behaviour. (a)-(d) each have a single aspect reinserted into the bare model; reinsertion is not sequential. (a) Upper curve has cell to cell exchange reinserted in the model with $\rho_{efN} = 1.8625 \times 10^{22}$ m^{-3}. (b) Upper curve has high losses removed from the model with $\rho_{efN} = 1.825 \times 10^{22}$ m^{-3}. (c) Upper curve has division reinserted into the model with $\rho_{efN} = 1.875 \times 10^{22}$ m^{-3}. (d) Upper curve has cell growth reinserted into the model with $\rho_{efN} = 2.075 \times 10^{22}$ m^{-3}.

In Figure 6.7(a) we show the cluster size distribution for both the bare percolation

model and the bare model with exchange between cells reinserted into the model. The external fixed nitrogen concentration had to be slightly adjusted to find power law behaviour with the exchange present, but the power law remains.

In Figure 6.7(b) we show the cluster size distribution for both the bare percolation model and the bare model with the large losses removed from the model. The external fixed nitrogen concentration had to be slightly adjusted to find power law behaviour with the lower losses, but the power law remains.

In Figure 6.7(c) we show the cluster size distribution for both the bare percolation model and the bare model with cell division reinserted into the model. This addition does not change the power law and the external fixed nitrogen concentration was not adjusted.

In Figure 6.7(d) we show the cluster size distribution for both the bare percolation model and the bare model with cell growth reinserted. To get the largest clusters to form, ρ_{efN} had to be significantly adjusted, from 1.875×10^{22} m^{-3} to 2.125×10^{22} m^{-3}. This adjustment is due to import of the external fixed nitrogen being proportional to length (see Equation 4.1). The cluster size distribution for the bare model plus cell growth has a significant difference from a power law with the cell growth, as it has a dip and a rise between cluster sizes of 10^2 and 10^3. Given that it is the only difference from the stripped down model, this feature is related to the changing cell size. The feature is not removed by adjusting ρ_{efN}, the only free parameter, except when it is able to bring the cutoff to a cluster size that is smaller than the location of the dip. The feature is similar to the deviation of the full model cluster size distribution from power law behaviour that is seen in Figure 6.3(a). The deviation from a power law in the full model is likely a smeared out version of the feature.

Figure 6.8 demonstrates how this dip and bump feature can be moved by changing the growth rate of the filament. A slower growth rate (longer cell period) increases the cluster size at which the feature occurs. This increase appears to occur linearly. This relationship between the plateau feature and the cell period was predicted by the temporal starvation distribution in Figure 6.5.

To achieve percolation the consumption rate was changed from being chosen every cell cycle, to being chosen every 10 seconds. Reinserting the choosing of growth rates every cell cycle resulted in a distribution that was very different from a power law
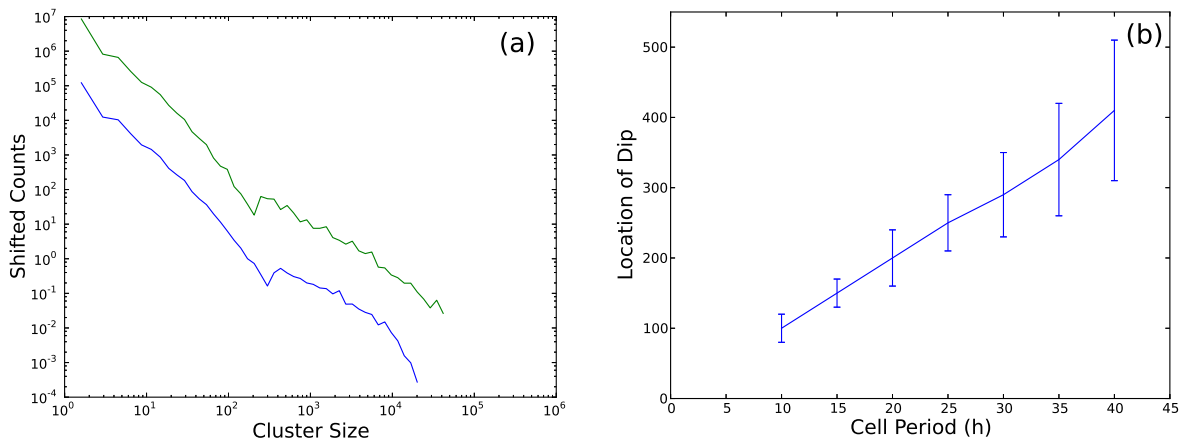
Figure 6.8: (a) Cluster size distribution in log-log format, using the same parameters as Figure 6.1. Systems have the bare model with cell growth reinserted. Upper curve is the same data as the upper curve in Figure 6.7(d). The lower curve is from a filament with half the growth rate of the upper curve, and highlights how the location of the dip moves. (b) is the location of the dip as the cell period is varied. Error bars represent the uncertainty in assigning the location of the dip.

(data not shown). This was because nothing about cells, including starvation status, could change until a cell cycle had passed.
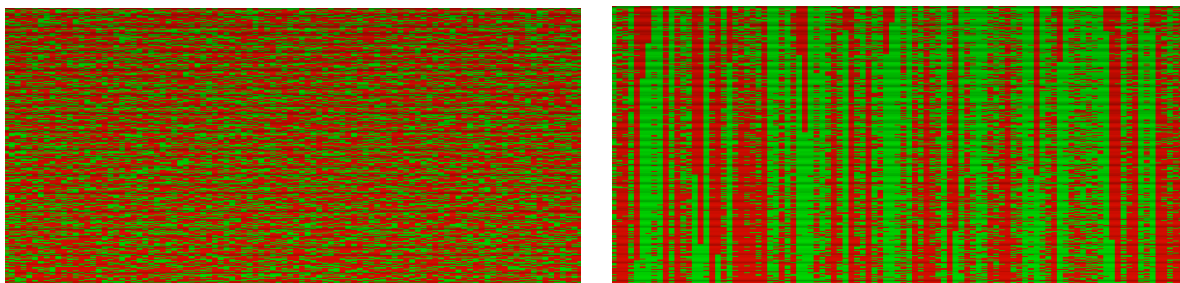
Figure 6.9: Spacetime visualizations of filament starvation, using the same parameters as Figure 6.1. Red represents starving cells and green represents non-starving cells. Each horizontal bar is the first 100 cells in the filament, and subsequent bars are separated by 1 minute. The left is for the bare model critical percolation system with the cluster distribution in Figure 6.6(a), as well as the bottom curve in Figures 6.7(a), (b), (c), and (d). The right is the top curve in Figure 6.7(d), with cell size dependent import of fixed nitrogen. This figure is meant to highlight the contrast in temporal correlations between the two spacetime visualizations.

In Figure 6.10 we have taken the full model of the filament and removed each of the elements of the model that cause deviations from power law behaviour separately

and together. Figure 6.10(a) shows the cluster size distribution for the full model again. Figure 6.10(b) is the cluster size distributions with the full model except with the cell size set to a constant for the calculation of fixed nitrogen import. At no external fixed nitrogen concentration do we see power law behaviour; instead all of the distributions decay with increasing cluster size. Figure 6.10(c) is the cluster size distribution with the full model except with the cell growth rate chosen frequently instead of every cell cycle. This does not yield power law behaviour. Figure 6.10(d) shows the cluster size distribution for the full model except with both of these changes (cell size set to a constant for import and frequent choice of growth rate). This system does yield power law behaviour and appears similar to the percolating bare model.

These two elements of the model prevent percolation because they introduce correlations in time. Figure 6.9 shows spacetime visualizations for both the bare model to achieve correlation, and the bare model with cell growth reinserted. The lack of correlations in the former and strong temporal correlations in the latter are immediately evident upon inspection.

A fixed nitrogen consumption rate that is maintained for a cell cycle introduces temporal correlations in whether or not that cell is starving. Cell size dependent fixed nitrogen import also introduces temporal correlations because short cells are more likely to be starving than large cells. Cell size does not jump within the range, but instead linearly increases, and so temporal correlations in the starving status of the cell are caused by cell size dependent fixed nitrogen import.

6.4.4   Discussion

The previous three subsections have shown that with the full cyanobacterial filament model the starving cells are not strongly correlated in space but correlated in time.

The weak spatial correlations are shown in Section 6.4.1 by the power law for the one dimensional percolation. The lack of strong spatial correlations is despite the exchange of fixed nitrogen between cells.

Strong temporal correlations occur because cells maintain the same growth rate until they divide and because the import of fixed nitrogen is related to cell size. The evidence that the cyanobacterial cells maintain the same growth rate is the variable nitrogen incorporation observed by Popa *et al* [8] that we discussed in detail
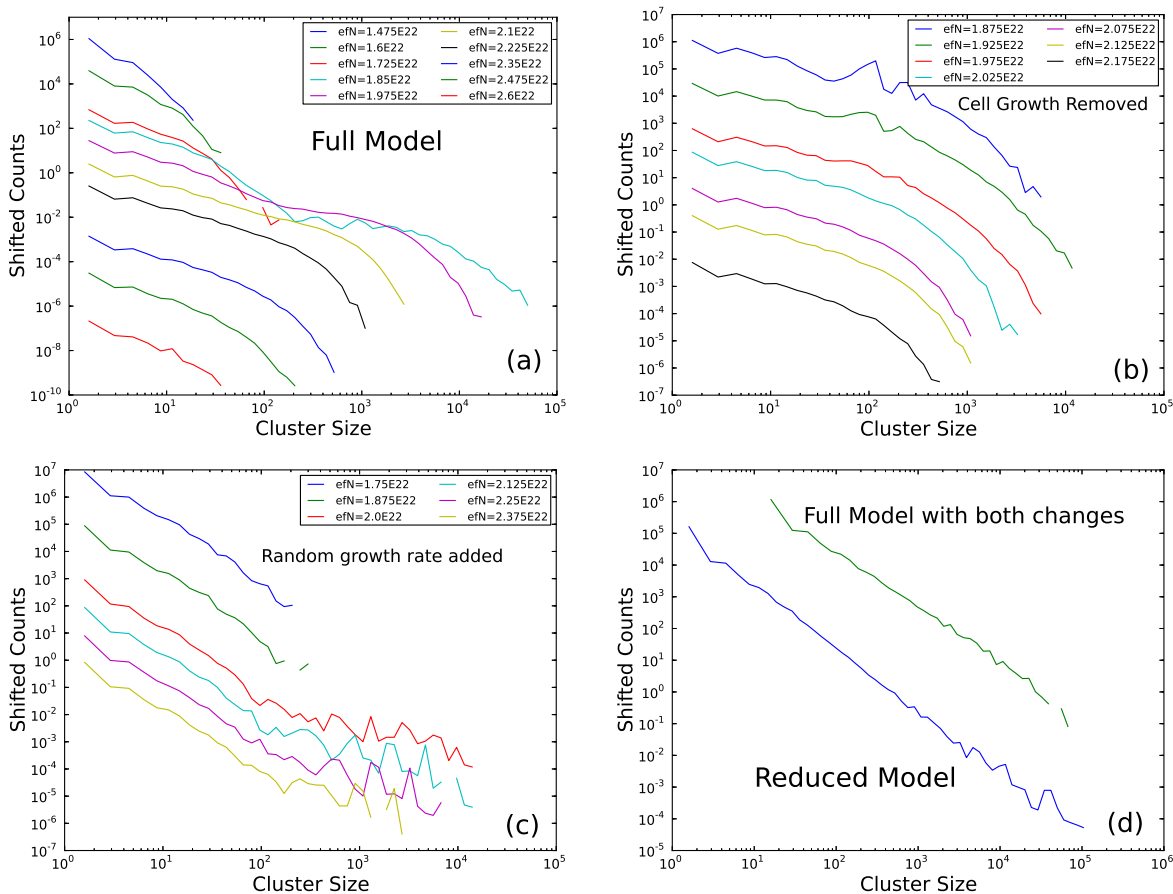
Figure 6.10: Cluster size distribution in log-log format, using the same parameters as Figure 6.1. (a) Full model. (b) Full model with fixed nitrogen import independent of cell size. (c) Full model with the growth rate of each cell chosen every 10 seconds. (d) Bare model reduced to percolation (the lower curve) and full model with frequent fixed nitrogen consumption rate changes and length independent import (the higher curve). $\rho_{efN} = 1.875 \times 10^{22}$ m^{-3} for the bare model and $\rho_{efN} = 1.85 \times 10^{22}$ m^{-3} for the full model with both changes. Data for full model with both changes is shifted to the right because of a factor of ten difference in the frequency of cluster sampling.

in Chapter 3. However, the variable nitrogen incorporation could be explained by each cell having a growth rate that changes in time, and each cell just has a different average growth rate (corresponding to the amount of fixed nitrogen incorporated) over that time. This is in contrast to our model, in which cells maintain a growth rate over a cell cycle, and choose a new one at division. The linear relationship between fixed nitrogen import and cell size is based upon the assumption that import is proportional to the surface area of the cell, and we do not have experimental data to

support it. It is therefore possible that both of the elements of the model that cause correlations and deviations from percolation behaviour could be plausibly omitted.

The dip in the cluster size distribution in Figure 6.7(d) seems to indicate a specific type of correlation in a two-dimensional system. The starving cells in this system have strong correlations in one direction, in which the starving cells connect for long stretches, and weak correlations in the other, in which the starving connections are relatively short and infrequent. The dip indicates the scale of these infrequent connections in the direction with weak correlations. It is possible that similar dips could suggest the presence of other scales in similar systems.

System growth is not usually considered in the percolation framework, but Figure 6.7(c) shows that cell division and growth of the system does not alter the percolating critical behaviour. The universality of the critical exponent describing the power law, despite cell division, is consistent with the expectation of the same percolating critical behaviour on irregular lattices as on regular lattices, and the observation that the rules of percolation can even be changed to some extent [137].

Correlation functions of fixed nitrogen concentration are calculated in Appendix A.

Chapter 7

Conclusion and Future Outlook

In this thesis we have developed quantitative models to investigate how filamentous, heterocystous cyanobacteria choose which cells will differentiate into heterocysts.

7.1   Summary of results

We have four main questions, each associated with an appropriate model.

The first question is how is the fixed nitrogen distributed and transported in the filament, and in Chapter 3 we examined the nitrogen distribution in the cyanobacterial filament between two heterocysts. This was motivated by two experimental distributions in the literature: that of Wolk *et al* [11] and that of Popa *et al* [8]. These distributions are quite different, with Wolk's nitrogen peaking at the heterocyst and dropping off away from the heterocyst, and Popa's nitrogen dipping significantly at the heterocyst and mostly being flat between the heterocysts. Wolk's distribution with its peak at the heterocyst agreed with our intuition, but Popa's was more recent with a better technique and better data. We found that both of these distributions are correct, and that the distribution of Wolk is for a much earlier time than the distribution of Popa. The initially counterintuitive distribution that Popa found is explained by the fact that vegetative cells grow and heterocysts do not. The vegetative cells accumulate the nitrogen when they grow, and this accumulation of nitrogen quickly dwarfs the magnitude of the freely diffusing fixed nitrogen that peaks at the heterocyst. We found that periplasmic transport of fixed nitrogen was not necessary to explain the experimental nitrogen distributions, but also could not be ruled out. Our model was successful when the cells had a maximum growth rate, instead of having growth linearly related to fixed nitrogen concentration at all fixed nitrogen levels, a situation which could not recover the Popa distribution.

This model formed the basis of the models that followed, though we removed periplasmic transport for simplicity and added rules for heterocyst placement.

The second question is why are the heterocysts patterned in the filament? Chapter 4 looked at the growth of the filament in the context of fixed nitrogen leakage and heterocyst placement strategies. We found that heterocyst frequency can optimize growth by balancing the fixed nitrogen produced by the heterocysts with the lack of growth that comes from heterocysts not dividing. By comparing the growth of the filament using different strategies for heterocyst placement we found that only with leakage present did the different strategies yield different growth rates. A local placement strategy was found to produce filaments that grew faster than filaments placing heterocysts randomly. For all three strategies, the heterocyst frequency was found to decrease continuously as the external fixed nitrogen level was increased. Heterocyst patterns for the random heterocyst placement strategies were significantly different than the experimental wild type patterns, but the pattern for the local placement strategy was qualitatively very similar.

This model, with small but nonzero leakage, and with an adaptive placement strategy was then used as the basis for the next model. In that case, we added fixed nitrogen storage and lateral inhibition. This produced a more refined model of heterocyst commitment.

We wanted to know how a filament recently deprived of fixed nitrogen breaks the starvation symmetry and decides which cells will become heterocysts. In Chapter 5 we recover many of the aspects of initial heterocyst differentiation upon nitrogen deprivation using a storage based commitment model. The model has a cell commit to heterocyst differentiation once the stored fixed nitrogen in the cell has decreased to a certain level. Cells that have been committed for a period of time then become mature heterocysts that produce fixed nitrogen. Committed cells and mature heterocysts inhibit nearby cells from commitment. This model qualitatively reproduced filament growth, heterocyst commitment, heterocyst pattern, and heterocyst frequency. It also produced a cell cycle effect on commitment and, by removing certain elements of the model, we qualitatively recovered mutant heterocyst patterns.

The final question is more curiosity driven: how are starving cells arranged in space and time in the absence of heterocysts? In Chapter 6 we showed that without heterocysts the quantitative model of the cyanobacterial filament was similar to the one dimensional percolation of starving cells and a more limited model exhibited two

dimensional percolation of starving cells. The elements of the model that caused deviations from percolation behaviour were identified.

7.2 Future outlook

There are some things that can be investigated within the fixed nitrogen model we have built without adding other elements such as other metabolites or genetic response.

In chapter 4 we investigated the steady state optimal heterocyst frequency for a local heterocyst placement strategy while varying the external fixed nitrogen concentration. A similar investigation could be performed with the more sophisticated heterocyst differentiation model of Chapter 5 to see what changes. In Chapter 5 the external fixed nitrogen concentration was held at zero after dropping from a high concentration. Instead of dropping to zero, the concentration could be dropped to the low but finite concentrations investigated in Chapter 4. The heterocyst frequency 24 hours after the drop in the fixed nitrogen concentration could be determined as the external fixed nitrogen concentration after the drop is varied. This would allow a similar plot to Figure 4.7 to be determined in this new situation. The heterocyst frequency after 24 hours of nitrogen deprivation in Chapter 5 is only after one generation of heterocyst differentiation and is largely controlled by the range of lateral inhibition. The steady state heterocyst frequency in Chapter 4 is after many generations of heterocyst differentiation and is controlled by the fixed nitrogen produced by the heterocysts. Because of this difference, the heterocyst frequency 24 hours after a drop in the external fixed nitrogen concentration might qualitatively vary differently with the external fixed nitrogen concentration after the drop compared to Figure 4.7, which varies smoothly. In particular, I think it is possible that the frequency could experience a shape similar to a step function as the external fixed nitrogen level after the drop is increased. This is because the storage has to run out for heterocyst commitment to occur, and the more fixed nitrogen the cells are able to import, the longer they can put this off, and eventually the earliest completion of differentiation will fall after the 24 hour mark.

In chapter 5 the storage fraction was set at a value that was determined from the literature. This value could be set to zero and the same type of measurements made

to see what the heterocyst pattern, commitment, etc. look like in this situation. One could also vary the amount of fixed nitrogen per unit length, g, as well as the amount of storage, and look at growth or whether the optimal inhibition range changes.

Cyanobacteria in the wild experience the day and night cycle, although those in the lab often do not. The model could be used to investigate the growth of the filament and the heterocyst pattern in a day and night cycle. Nitrogen fixation and growth would have to be varied depending on the light, and a full treatment might need to include carbohydrates. This type of investigation could help understand phenomena that the current canonical experiment misses.

In terms of understanding why a particular cell becomes a heterocyst and another does not, I think that our model (the most sophisticated version, from Chapter 5), which does everything with only fixed nitrogen, is probably at or near its limit. Our model is good at obtaining accurate overall distributions and other measurements for the entire filament, but we have approximated the commitment process. To progress further the model needs to include genes and proteins. Detailed modelling of the three genes that seem to be important to commitment, *hetR*, *patS*, and *patA*, will probably increase understanding. The goal of this type of work would not be to provide more accurate heterocyst spacing distributions and other measurements, although it probably would. It instead would work towards being able to accurately predict which individual cells were going to differentiate out of a cluster of starving cells. Certain information will have be gathered first. Do larger cells produce some or all of these proteins faster? Does the starvation state affect the rate of protein production? Also, the different ways PatS is hypothesized to act will all have to be investigated.

We have built into our model the resistance of the proheterocyst to its own PatS or HetN. In vivo this resistance is the result of genetic interactions, probably involving PatA, which are not completely understood. Inspiration in understanding how PatS acts and how it does not inhibit the cell that produced it may come from the Delta Notch system [138]. This system allows for the patterning of cells by the mutual inactivation of neighbouring cells and amplifies a small difference in signalling between two cells. It has also been modelled [139].

Bibliography

[1] Fogg G E. Growth and heterocyst production in *Anabaena cylindrica* Lemm. *New Phytol*, 43:164–175, 1944.

[2] Fogg G E. Growth and heterocyst production in *Anabaena cylindrica* Lemm. II. In relation to carbon and nitrogen metabolism. *Ann of Bot*, 13:241–259, 1949.

[3] Golden J W and Yoon H. Heterocyst development in *Anabaena*. *Curr Opin Microbiol*, 6:557–563, 2003.

[4] Cross M C and Hohenberg P C. Pattern formation outside of equilibrium. *Rev Mod Phys*, 65:851, 1993.

[5] Koch A J and Meinhardt H. Biological pattern formation: from basic mechanisms to complex structures. *Rev Mod Phys*, 66:1481–1507, 1994.

[6] Yoon H and Golden J W. Heterocyst pattern formation controlled by a diffusible peptide. *Science*, 282:935–938, 1998.

[7] Mullineaux C W, Mariscal V, Nenninger A, Khanum H, Herrero A, Flores E, and Adams D G. Mechanism of intercellular molecular exchange in heterocyst-forming cyanobacteria. *EMBO Journal*, 27:1299–1308, 2008.

[8] Popa R, Weber P K, Pett-Ridge J, Finzi J A, Fallon S J, Hutcheon I D, Nealson K H, and Capone D G. Carbon and nitrogen fixation and metabolite exchange in and between individual cells of *Anabaena oscillarioides*. *ISME Journal*, 1:354–360, 2007.

[9] Plischke M and Bergersen B. *Equilibrium statistical physics*. World Scientific Publishing Co, 2006.

[10] Ephrussi A and St Johnston D. Seeing is believing: the bicoid morphogen gradient matures. *Cell*, 116:143–152, 2004.

[11] Wolk C P, Austin S M, Bortins J, and Galonsky A. Autoradiographic localization of ^{13}N after fixation of ^{13}N-labeled nitrogen gas by a heterocyst-forming blue-green alga. *J Cell Biol*, 61:440–453, 1974.

[12] Adams D G and Duggan P S. Heterocyst and akinete differentiation in cyanobacteria. *New Phytol*, 144:3–33, 1999.

[13] Schopf J W. Disparate rates, differing fates: tempo and mode of evolution changed from the Precambrian to the Phanerozoic. *Proc Natl Acad Sci*, 91:6735–6742, 1994.

[14] Schopf J W. Are the oldest fossils cyanobacteria? In Roberts D M, Sharp P, Alderson G, and Collins M, editors, *Evolution of Microbial Life*, pages 23–61. Cambridge University Press, 1996.

[15] Flores E and Herrero A. Compartmentalized function through cell differentiation in filamentous cyanobacteria. *Nat Rev Microbiol*, 8:39–50, 2010.

[16] Kopp R E, Kirschvink J L, Hilburn I A, and Nash C Z. The paleoproterozoic snowball earth: a climate disaster triggered by the evolution of oxygenic photosynthesis. *Proc Nat Acad Sci*, 102:11131–11136, 2005.

[17] Falkowski P G. Tracing oxygen's imprint on earth's metabolic evolution. *Science*, 311:1724–1725, 2006.

[18] Fay P. *The Blue-Greens*. Edward Arnold Publishers Ltd, 1983.

[19] Herrero A, Muro-Pastor A M, Valladares A, and Flores E. Cellular differentiation and the NtcA transcription factor in filamentous cyanobacteria. *FEMS Microbiol Rev*, 28:469–487, 2004.

[20] Gallon J R. Reconciling the incompatible: N_2 fixation and O_2. *New Phytol*, 122:571–609, 1992.

[21] Fay P. Oxygen relations of nitrogen fixation in cyanobacteria. *Microbiol Rev*, 56:340–373, 1992.

[22] Bergman B, Gallon J R, Rai A N, and Stal L J. N_2 fixation by non-heterocystous cyanobacteria. *FEMS Microbiol Rev*, 19:139–185, 1997.

[23] Kasting J F and Siefert J L. Life and the evolution of earth's atmosphere. *Science*, 296:1066–1068, 2002.

[24] Giovannoni S J, Turner S, Olsen G F, Barns S, Lane D J, and Pace N R. Evolutionary relationships among cyanobacteria and green chloroplasts. *J Bacteriol*, 170:3584–3592, 1988.

[25] Wolk C P, Ernst A, and Elhai J. Heterocyst metabolism and development. In Bryant D A, editor, *The molecular biology of cyanobacteria*, pages 770–823. Kluwer Academic Publishers, 1994.

[26] Yoon H and Golden J W. PatS and products of nitrogen fixation control heterocyst pattern. *J Bacteriol*, 183:2605–2613, 2001.

[27] Ehira S, Ohmori M, and Sato N. Genome-wide expression analysis of the responses to nitrogen deprivation in the heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120. *DNA Res*, 10:97–113, 2003.

[28] Thiel T and Pratte B. Effect on heterocyst differentiation of nitrogen fixation in vegetative cells of the cyanobacterium *Anabaena variabilis* ATCC 29413. *J Bacteriol*, 183:280–286, 2001.

[29] Elder R G and Parker M. Growth response of a nitrogen fixer (*Anabaena flos-aquae*, cyanophyceae) to low nitrate. *J Phycol*, 20:296–301, 1984.

[30] Lynn M E, Bantle J A, and Ownby J D. Estimation of gene expression in heterocysts of *Anabaena variabilis* by using DNA-RNA hybridisation. *J Bacteriol*, 167:940–946, 1986.

[31] Meeks J C and Elhai J. Regulation of cellular differentiation in filamentous cyanobacteria in free-living and plant-associated symbiotic growth states. *Microbiol Mol Biol Rev*, 66:94–121, 2002.

[32] Walsby A E. The permeability of heterocysts to the gases nitrogen and oxygen. *Proc R Soc Lond B*, 226:345–366, 1985.

[33] Tel-Or E and Stewart W D P. Photosynthetic components and activities of nitrogen fixing, isolated heterocysts of *Anabaena cylindrica*. *Proc R Soc Lond B*, 198:61–86, 1977.

[34] Almon H and Bohme H. Components and activity of the photosynthetic electron transport system of intact heterocysts isolated from the blue-green alga *Nostoc muscorum*. *Biochim Biophys Acta*, 592:113–120, 1980.

[35] Fay P, Stewart W D P, Walsby A E, and Fogg G E. Is the heterocyst the site of nitrogen fixation in the blue-green algae? *Nature*, 220:810–812, 1968.

[36] Peterson R B and Wolk C P. Localization of an uptake hydrogenase in *Anabaena*. *Plant Physiol*, 61:688–691, 1978.

[37] Murry M A, Horne A J, and Benemann J R. Immunochemical evidence that nitrogenase is restricted to the heterocysts in *Anabaena cylindrica*. *Arch Microbiol*, 137:194–199, 1984.

[38] Bergman B, Lindblad P, and Rai A N. Nitrogenase in free-living and symbiotic cyanobacteria: immunoelectron microscopic localization. *FEMS Microbiol*, 35:75–78, 1986.

[39] Braun-Howland E B, Lindblad P, Nierzwicki-Bauer S A, and Bergman B. Dinitrogenase reductase (Fe-protein) of nitrogenase in the cyanobacterial symbionts of three *Azolla* species: localization and sequence of appearance during heterocyst differentiation. *Planta*, 176:319–332, 1988.

[40] Rai A N, Borthakur M, Singh S, and Bergman B. *Anthoceros-Nostoc* symbiosis: immunoelectronmicroscopic localization of nitrogenase, glutamine synthetase, phycoerythrin and ribulose-l,5-bisphosphate carboxylase/oxygenase in

the cyanobiont and the cultured (free-living) isolate *Nostoc* 7801. *J Gen Microbiol*, 135:385–395, 1989.

[41] Bergman B and Rai A. The *Nostoc-Nephroma* symbiosis: localization, distribution pattern and levels of key proteins involved in nitrogen and carbon metabolism of the cyanobiont. *Physiol Plant*, 77:216–224, 1989.

[42] Golden J W, Whorff L L, and Wiest D R. Independent regulation of *nifHDK* operon transcription and DNA rearrangement during heterocyst differentiation in the cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol*, 173:7098–7105, 1991.

[43] Wolk C P, Thomas J, and Shaffer P W. Pathway of nitrogen metabolism after fixation of ^{13}N-labeled nitrogen gas by the cyanobacterium, *Anabaena cylindrica*. *J Biol Chem*, 251:5027–5034, 1976.

[44] Wolk C P. Heterocysts. In Carr N G and Whitton B A, editors, *The Biology of Cyanobacteria*, pages 359–386. Blackwell Scientific Publishers, 1982.

[45] Thomas J, Meeks J C, Wolk C P, Shaffer P W, Austin S M, and Chien W S. Formation of glutamine from ^{13}N ammonia, ^{13}N dinitrogen, and ^{14}C glutamate by heterocysts isolated from *Anabaena cylindrica*. *J Bacteriol*, 129:1545–1555, 1977.

[46] Toyoshima M, Sasaki N V, Fujiwara M, Ehira S, Ohmori M, and Sato N. Early candidacy for differentiation into heterocysts in the filamentous cyanobacterium *Anabaena* sp. PCC 7120. *Arch Microbiol*, 192:23–31, 2010.

[47] Khudyakov I Y and Golden J W. Different functions of HetR, a master regulator of heterocyst differentiation in *Anabaena* sp. PCC 7120, can be separated by mutation. *Proc Natl Acad Sci*, 101:16040–16045, 2004.

[48] Bradley S and Carr N G. Heterocyst and nitrogenase development in *Anabaena cylindrica*. *J Gen Microbiol*, 96:175–184, 1976.

[49] Adams D G. Heterocyst formation in cyanobacteria. *Curr Opin Microbiol*, 3:618–624, 2000.

[50] Neilson A, Rippka R, and Kunisawa R. Heterocyst formation and nitrogenase synthesis in *Anabaena* sp. *Arch Microbiol*, 76:139–150, 1971.

[51] Orozco C C, Risser D D, and Callahan S M. Epistasis analysis of four genes from *Anabaena* sp. strain PCC 7120 suggests a connection between PatA and PatS in heterocyst pattern formation. *J Bacteriol*, 188:1808–1816, 2006.

[52] Grossman A R, Schaefer M R, Chiang G G, and Collier J L. The responses of cyanobacteria to environmental conditions: light and nutrients. In Bryant D A, editor, *The molecular biology of cyanobacteria*, pages 641–676. Kluwer Academic Publishers, 1994.

[53] Asai H, Iwamori S, Kawai K, Ehira S, Ishihara J, Aihara K, Shoji S, and Iwasaki H. Cyanobacterial cell lineage analysis of the spatiotemporal *hetR* expression profile during heterocyst pattern formation in *Anabaena* sp. PCC 7120. *PLOS One*, 4:e7371, 2009.

[54] Buikema W J and Haselkorn R. Characterization of a gene controlling heterocyst differentiation in the cyanobacterium *Anabaena* 7120. *Genes Dev*, 5:321–330, 1991.

[55] Black T A, Cai Y, and Wolk C P. Spatial expression and autoregulation of *hetR*, a gene involved in the control of heterocyst development in *Anabaena*. *Mol Microbiol*, 9:77–84, 1993.

[56] Allard J F, Hill A L, and Rutenberg A D. Heterocyst patterns without patterning proteins in cyanobacterial filaments. *Dev Biol*, 312:427–434, 2007.

[57] Liang J, Scappino L, and Haselkorn R. The *patA* gene product, which contains a region similar to CheY of *Escherichia coli*, controls heterocyst pattern formation in the cyanobacterium *Anabaena* 7120. *Proc Natl Acad Sci*, 89:5655–5659, 1992.

[58] Maldener I, Lockau W, Cai Y, and Wolk C P. Calcium-dependent protease of the cyanobacterium *Anabaena*: molecular cloning and expression of the gene in *Escherichia coli*, sequencing and site-directed mutagenesis. *Mol Gen Genet*, 225:113–120, 1991.

[59] Bohme H and Haselkorn R. Molecular cloning and nucleotide sequence analysis of the gene coding for heterocyst ferredoxin from the cyanobacterium *Anabaena* sp. strain PCC 7120. *Mol Gen Genet*, 214:278–285, 1988.

[60] Pernil R, Picossi S, Mariscal V, Herrero A, and Flores E. ABC-type amino acid uptake transporters Bgt and N-II of *Anabaena* sp strain PCC 7120 share an ATPase subunit and are expressed in vegetative cells and heterocysts. *Mol Micro*, 67:1067–1080, 2008.

[61] Borthakur P B, Orozco C C, Young-Robbins S S, Haselkorn R, and Callahan S M. Inactivation of *patS* and *hetN* causes lethal levels of heterocyst differentiation in the filamentou cyanobacterium *Anabaena* sp. PCC 7120. *Mol Microbiol*, 57:111–123, 2005.

[62] Callahan S M and Buikema W J. The role of HetN in maintenance of the heterocyst pattern in *Anabaena* sp. PCC 7120. *Mol Micro*, 40:941–950, 2001.

[63] Bauer C C, Ramaswamy K S, Endley S, Scappino L A, Golden J W, and Haselkorn R. Suppression of heterocyst differentiation in *Anabaena* PCC 7120 by a cosmid carrying wild-type genes encoding enzyme for fatty acid synthesis. *FEMS Microbiol Lett*, 151:23–30, 1997.

[64] Li H, Sherman D M, Bao S, and Sherman L A. Pattern of cyanophycin accumulation in nitrogen-fixing and non-nitrogen-fixing cyanobacteria. *Arch Microbiol*, 176:9–18, 2001.

[65] Picossi S, Valladares A, Flores E, and Herrero A. Nitrogen-regulated genes for the metabolism of cyanophycin, a bacterial nitrogen reserve polymer. *J Biol Chem*, 279:11582–11592, 2004.

[66] Ziegler K, Stephan D P, Pistorius E K, Ruppel H G, and Lockau W. A mutant of the cyanobacterium *Anabaena variabilis* ATCC 29413 lacking cyanophycin sythetase: growth properties and ultrastructural aspects. *FEMS Microbiol Lett*, 196:13–18, 2001.

[67] Allen M M. Cyanobacterial cell inclusions. *Ann Rev Microbiol*, 38:1–25, 1984.

[68] Simon R D. Measurement of the cyanophycin granule polypeptide contained in the blue-green alga *Anabaena cylindrica*. *J Bacteriol*, 114:1213–1216, 1973.

[69] Buikema W J and Haselkorn R. Molecular genetics of cyanobacterial development. *Annu Rev Plant Physiol Plant Mol Biol*, 44:33–52, 1993.

[70] Baier K, Lehmann H, Stephan D P, and Lockau W. NblA is essential for phycobilisome degradation in *Anabaena* sp. strain PCC 7120 but not for development of functional heterocysts. *Microbiology*, 150:2739–2749, 2004.

[71] Turing A M. The chemical basis of morphogenesis. *Philos Trans R Soc London Ser B*, 237:37–72, 1952.

[72] Mitchison G J, Wilcox M, and Smith R J. Measurement of an inhibitory zone. *Science*, 191:866–868, 1976.

[73] Adams D G and Carr N G. Control of heterocyst development in the cyanobacterium *Anabaena cylindrica*. *J Gen Microbiol*, 135:839–849, 1989.

[74] Wolk C P and Quine M P. Formation of one-dimensional patterns by stochastic processes and by filamentous blue-green algae. *Dev Biol*, 46:370–382, 1975.

[75] Brown A I and Rutenberg A D. Reconciling cyanobacterial fixed-nitrogen distributions and transport experiments with quantitative modelling. *Phys Biol*, 9:016007, 2012.

[76] Haselkorn R. Cell-cell communication in filamentous cyanobacteria. *Molecular Microbiology*, 70:783–785, 2008.

[77] Giddings T H and Staehelin L A. Plasma membrane architecture of *Anabaena cylindrica*: occurrence of microplasmodesmata and changes associated with heterocyst development and the cell cycle. *Cytobiologie*, 16:235–249, 1978.

[78] Giddings T H and Staehelin L A. Observation of microplasmodesmata in both heterocyst-forming and non-heterocyst forming filamentous cyanobacteria by freeze-fracture electron microscopy. *Arch Microbiol*, 129:295–298, 1981.

[79] Lang N J and Fay P. The heterocysts of blue-green algae. I. Ultrastructural integrity after isolation. *Proc R Soc Lond B*, 178:193–203, 1971.

[80] Mariscal V, Herrero A, Nenninger A, Mullineaux C W, and Flores E. Functional dissection of the three-domain SepJ protein joining the cells in cyanobacterial trichomes. *Mol Micro*, 79:1077–1088, 2011.

[81] Flores E, Pernil R, Muro-Pastor A M, Mariscal V, Maldener I, Lechno-Yossef S, Fan Q, Wolk C P, and Herrero A. Septum-localized protein required for filament integrity and diazotrophy in the heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol*, 189:3884–3890, 2007.

[82] Flores E, Herrero A, Wolk C P, and Maldener I. Is the periplasm continuous in filamentous multicellular cyanobacteria? *Trends Microbiol*, 14:439–443, 2006.

[83] Mariscal V, Herrero A, and Flores E. Continuous periplasm in a filamentous, heterocyst-forming cyanobacterium. *Mol Microbiol*, 65:1139–1145, 2007.

[84] Kumar K, Mella-Herrera R A, and Golden J W. Cyanobacterial heterocysts. *Cold Spring Harb Perspect Biol*, 2:a000315, 2010.

[85] Zhang L-C, Chen Y-F, Chen W-L, and Zhang C-C. Existence of periplasmic barriers preventing green fluorescent protein diffusion from cell to cell in the cyanobacterium *Anabaena* sp. strain PCC 7120. *Mol Micro*, 70:814–823, 2008.

[86] Nicolaisen K, Mariscal V, Bredemeier R, Pernil R, Moslavac S, López-Igual R, Maldener I, Herrero A, Schlieff E, and Flores E. The outer membrane of a heterocyst-forming cyanobacterium is a permeability barrier for uptake of metabolites that are exchanged between cells. *Mol Micro*, 74:58–70, 2009.

[87] Montesinos M L, Herrero A, and Flores E. Amino acid transport systems required for diazotrophic growth in the cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol*, 177:3150–3157, 1995.

[88] Picossi S, Montesinos M L, Pernil R, Lichtle C, Herrero A, and Flores E. ABC-type neutral amino acid permease N-I is required for optimal diazotrophic growth and is repressed in the heterocysts of *Anabaena* sp. strain PCC 7120. *Mol Microbiol*, 57:1582–1592, 2005.

[89] Nenninger A, Mastroianni G, and Mullineaux C W. Size dependence of protein diffusion in the cytoplasm of *Escherichia coli*. *J Bacteriol*, 192:4535–4540, 2010.

[90] Mullineaux C W, Nenninger A, Ray N, and Robinson C. Diffusion of green fluorescent protein in three cell environments in *Escherichia Coli*. *J Bacteriol*, 188:3442–3448, 2006.

[91] Rees D C, Johnson E, and Lewinson O. ABC transporters: the power to change. *Nat Rev Mol Cell Biol*, 10:218–227, 2009.

[92] Philips R, Kondev J, and Theriot J. *Physical Biology of the Cell*. Garland Science, New York, 2009.

[93] Paerl H W. Role of heterotrophic bacteria in promoting N_2 fixation by *Anabaena* in aquatic habitats. *Microb Ecol*, 4:215–231, 1978.

[94] Thiel T. Protein turnover and heterocyst differentiation in the cyanobacterium *Anabaena variabilis*. *J Phycol*, 26:50–54, 1990.

[95] Herrero A and Flores E. Transport of basic amino acids by the dinitrogen-fixing cyanobacterium *Anabaena* PCC 7120. *J Biol Chem*, 265:3931–3935, 1990.

[96] Powell E O. Growth rate and generation time of bacteria, with special reference to continuous culture. *J Gen Microbiol*, 15:492–511, 1956.

[97] Dunn J H and C P Wolk. Composition of the cellular envelopes of *Anabaena cylindrica*. *J Bacteriol*, 103:153–158, 1970.

[98] Cobb H D and Myers J. Comparative studies of nitrogen fixation and photosynthesis in *Anabaena cylindrica*. *Amer Jour Bot*, 51:753–762, 1964.

[99] Flores E and Herrero A. Assimilatory nitrogen metabolism and its regulation. In Bryant D A, editor, *The molecular biology of cyanobacteria*, pages 487–517. Kluwer Academic Publishers, 1994.

[100] Kovarova-Kovar D and Egli T. Growth kinetics of suspended microbial cell: from single substrate-controlled growth to mixed-substrate kinetics. *Microbiol Mol Biol Rev*, 62:646–666, 1998.

[101] Ikeda T P, Shauger A E, and Kustu S. *Salmonella typhimurium* apparently perceives external nitrogen limitation as internal glutamine limitation. *J Mol Biol*, 259:589–607, 1996.

[102] Brauer M J, Yuan J, Bennett B D, Lu W, Kimball E, Botstein D, and Rabinowitz J D. Conservation of the metabolomic response to starvation across two divergent microbes. *P Natl Acad Sci (USA)*, 103:19302–19307, 2006.

[103] Hart Y, Madar D, Yuan J, Bren A, Mayo A E, Rabinowitz J D, and Alon U. Robust control of nitrogen assimilation by a bifunctional enzyme in *E. coli*. *Mol Cell*, 41:117–127, 2011.

[104] Fehr M, Ehrhardt D W, Lalonde S, and Frommer W B. Minimally invasive dynamic imaging of ions and metabolites in living cells. *Curr Opin Plant Biol*, 7:345–351, 2004.

[105] Brown A I and Rutenberg A D. Heterocyst placement strategies to maximize the growth of cyanobacterial filaments. *Phys Biol*, 9:046002, 2012.

[106] Jones K M, Buikema W J, and Haselkorn R. Heterocyst-specific expression of *patB*, a gene required for nitrogen fixation in *Anabaena* sp. strain PCC 7120. *J Bacteriol*, 185:2306–2314, 2003.

[107] Walsby A E and Fogg G E. The extracellular products of *Anabaena cylindrica* lemm. III. excretion and uptake of fixed nitrogen. *Br Phycol J*, 10:339–345, 1975.

[108] Gibson C E and Smith R V. Freshwater plankton. In Carr N G and Whitton B A, editors, *The biology of cyanobacteria*, pages 492–513. University of California Press, 1982.

[109] Wetzel R G. *Limnology*. W. B. Saunders Company, Toronto, 1975.

[110] Lund J W G. The ecology of the freshwater phytoplankton. *Biol Rev*, 40:231–293, 1965.

[111] Nielsen E S. Productivity of the oceans. *Annu Rev Pl Physiol*, 11:341–362, 1960.

[112] Reynolds C S and Walsby A E. Water-blooms. *Biol Rev*, 50:437–481, 1975.

[113] Stumm W and Morgan J J. *Aquatic chemistry*. John Wiley and Sons, New York, 1996.

[114] Downing J A and McCauley E. The nitrogen:phosphorus relationship in lakes. *Limnol Oceanogr*, 37:936–945, 1992.

[115] Charpy L, Palinska K A, Casareto B, Langlade M J, Suzuki Y, Abed R M M, and Golubic S. Dinitrogen-fixing cyanobacteria in microbial mats of two shallow coral reef ecosystems. *Microb Ecol*, 59:174–186, 2010.

[116] Stacey G, Baalen C V, and Tabita F R. Isolation and characterization of a marine *Anabaena* sp. capable of rapid growth on molecular nitrogen. *Arch Microbiol*, 114:197–201, 1977.

[117] Tyrrell T. The relative influences of nitrogen and phosphorus on oceanic primary production. *Nature*, 400:525–531, 1999.

[118] Ogawa R E and Carr J F. The influence of nitrogen on heterocyst production in blue-green algae. *Limnol Oceanogr*, 14:342–351, 1969.

[119] Horne A J, Sandusky J C, and Carmiggelt C J W. Nitrogen fixation in Clear Lake, California. 3. repetitive synoptic sampling of the spring *Aphanizomenon* blooms. *Limnol Oceanogr*, 24:316–328, 1979.

[120] Imhof M and Schlotterer C. Fitness effects of advantageous mutations in evolving *Escherichia coli* populations. *Proc Natl Acad Sci*, 98:1113–1117, 2000.

[121] Sultan S E. Phenotypic plasticity for plant development, function and life history. *Trends Plant Sci*, 5:537–542, 2000.

[122] Holtmaat A and Svoboda K. Experience-dependent structural synaptic plasticity in the mammalian brain. *Nat Rev Neurosci*, 10:647–658, 2009.

[123] Balaban N Q, Merrin J, Chait R, Kowalik L, and Leibler S. Bacterial persistence as a phenotypic switch. *Science*, 305:1622–1625, 2004.

[124] Bennett M R and Hasty J. Microfluidic devices for measuring gene network dynamics in single cells. *Nat Rev Genet*, 10:628–638, 2009.

[125] Turpin D H, Layzell D B, and Elrifi I R. Modeling the C economy of *Anabaena flos-aquae*. *Plant Physiol*, 78:746–752, 1985.

[126] Cumino A C, Marcozzi C, Barreiro R, and Salerno G L. Carbon cycling in *Anabaena* sp. PCC 7120. sucrose synthesis in the heterocysts and possible role in nitrogen fixation. *Plant Physiol*, 143:1385–1397, 2007.

[127] Bissett W P, Walsh J J, Dieterle D A, and Carder K L. Carbon cycling in the upper waters of the Sargasso sea: I. numerical simulation of differential carbon and nitrogen fluxes. *Deep-Sea Res*, 46:205–269, 1999.

[128] Neumann T. Towards a 3D-ecosystem model of the Baltic sea. *J Marine Syst*, 25:405–419, 2000.

[129] Sakr S, Jeanjean R, Zhang C, and Arcondeguy T. Inhibition of cell division suppresses heterocyst development in *Anabaena* sp. strain PCC 7120. *J Bacteriol*, 188:1396–1404, 2006.

[130] Bogorad L. Phycobiliproteins and complementary chromatic adaptation. *Ann Rev Plant Physiol*, 26:369–401, 1975.

[131] Wolk C P. Heterocyst formation. *Annu Rev Genet*, 30:59–78, 1996.

[132] Frias J E, Flores E, and Herrero A. Nitrate assimilation gene cluster from the heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol*, 179:477–486, 1997.

[133] Hebbar P B and Curtis S E. Characterization of *devH*, a gene encoding a putative DNA binding protein required for heterocyst function in *Anabaena* sp. strain PCC 7120. *J Bacteriol*, 182:3572–3581, 2000.

[134] Aldea M R, Mella-Herrera R A, and Golden J W. Sigma factor genes *sigC*, *sigE*, and *sigG* are upregulated in heterocysts of the cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol*, 189:8392–8396, 2007.

[135] Wei T F, Ramasubramanian T S, and Golden J W. *Anabaena* sp. strain PCC 7120 *ntcA* gene required for growth on nitrate and heterocyst development. *J Bacteriol*, 176:4473–4482, 1994.

[136] Stauffer D and Aharony A. *Introduction to percolation theory.* Taylor and Francis, 2004.

[137] Isichenko M B. Percolation, statistical topography, and transport in random media. *Rev Mod Phys*, 64:961–1043, 1992.

[138] Sprinzak D, Lakhanpal A, LeBon L, Santat L A, Fontes M E, Anderson G A, Garcia-Ojalvo J, and Elowitz M B. Cis-interactions between Notch and Delta generate mutually exclusive signalling states. *Nature*, 465:86–90, 2010.

[139] Sprinzak D, Lakhanpal A, LeBon L, Garcia-Ojalvo J, and Elowitz M B. Mutual inactivation of notch receptor and ligands facilitates developmental patterning. *PLOS Comput Biol*, 7:e1002069, 2011.

Appendix A

Correlation Functions

Spatial and temporal correlations are important to understanding the pattern of starving cells, and we wanted to investigate the correlations of fixed nitrogen in a filament with no heterocysts. The goal is to determine what sets the spatial and temporal scales in the perturbative limit and then check the generality of these results. We also seek the functional form of the correlations.

A.1   Calculation

We first model the change in fixed nitrogen density in time, $\dot{\rho}$, in a continuum with several terms. The first is a diffusion term, $D\nabla^2\rho$. The second is a consumption/uptake term G, which represents the balance of consumption of fixed nitrogen due to growth and the uptake of fixed nitrogen from outside the filament. The third is a leakage term proportional to the density, $-\rho/\tau$. The fourth term is the noise term, η, which will embody the stochasticity of consumption from stochastic growth:

$$\dot{\rho} = D\nabla^2\rho + G - \frac{\rho}{\tau} + \eta(x,t). \tag{A.1}$$

We will use the following Fourier Transform Pair conventions:

$$
\begin{aligned}
\rho(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde{\rho}(k)e^{ikx}dk, \\
\tilde{\rho}(k) &= \int_{-\infty}^{\infty}\rho(x)e^{-ikx}dx, \\
\rho(t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde{\rho}(\omega)e^{i\omega t}d\omega, \\
\tilde{\rho}(\omega) &= \int_{-\infty}^{\infty}\rho(t)e^{-i\omega t}dt.
\end{aligned}
\tag{A.2}
$$

We, conventionally, choose η to be white noise with no scales in time or space:

$$
\begin{aligned}
<\eta(x,t)> &= 0, \\
<\eta(x,t)\eta(x',t')> &= a\delta(x-x')\delta(t-t').
\end{aligned}
\tag{A.3}
$$

Solving the differential equation A.1 yields

$$i\omega\rho(k,\omega) = -Dk^2\rho(k,\omega) + G\delta(k)\delta(\omega) - \frac{1}{\tau}\rho(k,\omega) + \eta(k,\omega)$$

Solving for $\rho(k,\omega)$

$$\rho(k,\omega) = \frac{\eta(k,\omega) + G\delta(k)\delta(\omega)}{i\omega + Dk^2 + \frac{1}{\tau}}$$

The correlation function is then (terms with a single η do not survive because the average is zero)

$$< \rho(k,\omega)\rho(k',\omega') > = \left(\frac{1}{i\omega + Dk^2 + \frac{1}{\tau}}\right)\left(\frac{1}{i\omega' + Dk'^2 + \frac{1}{\tau}}\right)$$
$$[< \eta(k,\omega)\eta(k',\omega') > + G\delta(k)\delta(k')\delta(\omega)\delta(\omega')] \quad \text{(A.4)}$$

To find $< \eta(k,\omega)\eta(k',\omega') >$ we use the noise characteristics of Equation A.3

$$< \eta(k,\omega)\eta(k',\omega') > = \int_{-\infty}^{\infty}dx\int_{-\infty}^{\infty}dx'\int_{-\infty}^{\infty}dt\int_{-\infty}^{\infty}dt' < \eta(x,t)\eta(x',t') > e^{-ikx}e^{-ik'x'}e^{-i\omega t}e^{-i\omega't'}$$

$$= \int_{-\infty}^{\infty}dx\int_{-\infty}^{\infty}dx'\int_{-\infty}^{\infty}dt\int_{-\infty}^{\infty}dt'a\delta(x-x')\delta(t-t')e^{-ikx}e^{-ik'x'}e^{-i\omega t}e^{-i\omega't'}$$

$$= a\int_{-\infty}^{\infty}dx e^{-i(k+k')x}\int_{-\infty}^{\infty}dt e^{-i(\omega+\omega')t}$$

$$= a(2\pi)^2\delta(k+k')\delta(\omega+\omega') \quad \text{(A.5)}$$

Putting Equation A.5 back into Equation A.4 gives

$$< \rho(k,\omega)\rho(k',\omega') > = \left(\frac{1}{i\omega + Dk^2 + \frac{1}{\tau}}\right)\left(\frac{1}{i\omega' + Dk'^2 + \frac{1}{\tau}}\right)[a(2\pi)^2\delta(k+k')\delta(\omega+\omega')$$
$$+ G\delta(k)\delta(k')\delta(\omega)\delta(\omega')] \quad \text{(A.6)}$$

This expression will now be taken back to real space. There are two terms.

$$< \rho(x,t)\rho(x',t') > = < \rho(x,t)\rho(x',t') >_1 + < \rho(x,t)\rho(x',t') >_2$$

The first term is

$$< \rho(x,t)\rho(x',t') >_1 = \frac{1}{(2\pi)^4}\int_{-\infty}^{\infty}d\omega\int_{-\infty}^{\infty}d\omega'\int_{-\infty}^{\infty}dk\int_{-\infty}^{\infty}dk' < \rho(k,\omega)\rho(k',\omega') >_1 e^{ikx}e^{ik'x'}e^{i\omega t}e^{i\omega't'}$$

$$= \frac{a}{(2\pi)^2}\int_{-\infty}^{\infty}d\omega\int_{-\infty}^{\infty}d\omega'\int_{-\infty}^{\infty}dk\int_{-\infty}^{\infty}dk'\left(\frac{1}{i\omega + Dk^2 + \frac{1}{\tau}}\right)\left(\frac{1}{i\omega' + Dk'^2 + \frac{1}{\tau}}\right)$$
$$\delta(k+k')\delta(\omega+\omega')e^{ikx}e^{ik'x'}e^{i\omega t}e^{i\omega't'}$$

$$= \frac{a}{(2\pi)^2} \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} dk \left(\frac{1}{\omega^2 + (Dk^2 + \frac{1}{\tau})^2} \right) e^{ik(x-x')} e^{i\omega(t-t')} \tag{A.7}$$

The second term is

$$< \rho(x,t)\rho(x',t') >_2 = \frac{1}{(2\pi)^4} \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} d\omega' \int_{-\infty}^{\infty} dk \int_{-\infty}^{\infty} dk' < \rho(k,\omega)\rho(k',\omega') >_2 e^{ikx} e^{ik'x'} e^{i\omega t} e^{i\omega' t'}$$

$$\begin{aligned} = \quad & G \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} d\omega' \int_{-\infty}^{\infty} dk \int_{-\infty}^{\infty} dk' \left(\frac{1}{i\omega + Dk^2 + \frac{1}{\tau}} \right) \left(\frac{1}{i\omega' + Dk'^2 + \frac{1}{\tau}} \right) \\ & \delta(k)\delta(k')\delta(\omega)\delta(\omega') e^{ikx} e^{ik'x'} e^{i\omega t} e^{i\omega' t'} \\ = \quad & G\tau^2 \end{aligned} \tag{A.8}$$

Returning to the first term, Equation A.7, the ω integral is performed first. Looking at only the ω integral

$$\int_{-\infty}^{\infty} d\omega \left(\frac{1}{\omega^2 + (Dk^2 + \frac{1}{\tau})^2} \right) e^{i\omega(t-t')}$$

We have two poles where $\omega^2 + (Dk^2 + \frac{1}{\tau})^2 = 0$. These occur at $\omega = \pm i(Dk^2 + \frac{1}{\tau})$. Using two semicircular contours, each with the flat part on the real axis, only the positive pole is inside the upper contour. Using this upper pole to find the residue

$$R = \lim_{\omega \to i\left(Dk^2 + \frac{1}{\tau}\right)} \left(\omega - i\left(Dk^2 + \frac{1}{\tau}\right) \right) \left(\frac{1}{\omega^2 + \left(Dk^2 + \frac{1}{\tau}\right)^2} \right) e^{i\omega(t-t')}$$

$$= \frac{e^{-(Dk^2 + \frac{1}{\tau})(t-t')}}{2i(Dk^2 + \frac{1}{\tau})} \tag{A.9}$$

This means the integral

$$\int_{-\infty}^{\infty} d\omega \left(\frac{1}{\omega^2 + (Dk^2 + \frac{1}{\tau})^2} \right) e^{i\omega(t-t')} = 2\pi i R = \frac{\pi}{Dk^2 + \frac{1}{\tau}} e^{-(Dk^2 + \frac{1}{\tau})(t-t')} \tag{A.10}$$

This gives

$$< \rho(x,t)\rho(x',t') >_1 = \frac{a}{(2\pi)^2} \int_{-\infty}^{\infty} dk e^{ik(x-x')} \frac{\pi}{Dk^2 + \frac{1}{\tau}} e^{-(Dk^2 + \frac{1}{\tau})(t-t')}$$

Here we set $x' = 0$ and $t' = 0$, which simplifies the expression without loss of generality.

$$< \rho(x,t)\rho(0,0) >_1 = \frac{a}{4\pi} e^{-t/\tau} \int_{-\infty}^{\infty} \frac{dk}{Dk^2 + \frac{1}{\tau}} e^{ikx - Dk^2 t} \tag{A.11}$$

First we want to try two special cases. The first special case is $t = 0$. This gives us

$$< \rho(x,0)\rho(0,0) >_1 = \frac{a}{4\pi} \int_{-\infty}^{\infty} \frac{dk}{Dk^2 + \frac{1}{\tau}} e^{ikx}$$

Performing the contour integral yields

$$< \rho(x,0)\rho(0,0) >_1 = \frac{a}{4\pi} 2\pi i R = \frac{a}{4\pi} 2\pi i \frac{1}{2i} \sqrt{\frac{\tau}{D}} e^{-x/\sqrt{D\tau}} \tag{A.12}$$

$$< \rho(x,0)\rho(0,0) >_1 = \frac{a}{4} \sqrt{\frac{\tau}{D}} e^{-x/\sqrt{D\tau}} \tag{A.13}$$

With both terms the $t = 0$ case is

$$< \rho(x,0)\rho(0,0) >= \frac{a}{4} \sqrt{\frac{\tau}{D}} e^{-x/\sqrt{D\tau}} + G\tau^2 \tag{A.14}$$

Now we want to try the second special case: $x = 0$. From Equation A.11 this gives

$$< \rho(0,t)\rho(0,0) >_1 = \frac{a}{4\pi} e^{-t/\tau} \int_{-\infty}^{\infty} \frac{dk}{Dk^2 + \frac{1}{\tau}} e^{-Dk^2 t}$$

This can be rewritten as

$$< \rho(0,t)\rho(0,0) >_1 = \frac{a}{2\pi} e^{-t/\tau} \int_0^{\infty} \frac{t \cdot dk}{Dk^2 t + \frac{t}{\tau}} e^{-Dk^2 t} = \frac{at}{2\pi} e^{-t/\tau} \int_0^{\infty} \frac{dk'}{(k')^2 + \frac{t}{\tau}} e^{-(k')^2}$$

where $k' = \sqrt{Dt}k$. This is now in the form for (solution from Mathematica)

$$\int_0^{\infty} \frac{dk}{k^2 + a} e^{-k^2} = \frac{e^a \pi \cdot erfc(\sqrt{a})}{2\sqrt{a}}$$

Therefore

$$< \rho(0,t)\rho(0,0) >_1 = \frac{at}{2\pi} e^{-t/\tau} \cdot \frac{e^{t/\tau} \pi \cdot erfc(t/\tau)}{2t/\tau} = \frac{a\tau}{4} erfc(t/\tau) \tag{A.15}$$

With both terms the $x = 0$ case is

$$< \rho(0,t)\rho(0,0) >= \frac{a\tau}{4} erfc(t/\tau) + G\tau^2 \tag{A.16}$$

Now we want to go back to Equation A.11 and attempt generally. By completing the square the argument of the exponential can be changed to

$$ikx - Dk^2 t = -Dt \left[k^2 - \frac{ix}{Dt} k + \left(\frac{ix}{2Dt} \right)^2 \right] - \frac{x^2}{4Dt} = -Dt \left(k - \frac{ix}{2Dt} \right)^2 - \frac{x^2}{4Dt}$$

Putting this into A.11 yields

$$< \rho(x,t)\rho(0,0) >_1 = \frac{a}{4\pi}e^{-t/\tau}e^{-\frac{x^2}{4Dt}}\int_{-\infty}^{\infty}\frac{dk}{Dk^2+\frac{1}{\tau}}e^{-Dt\left(k-\frac{ix}{2Dt}\right)^2} \qquad (A.17)$$

Now we want to expand the $\frac{1}{Dk^2+\frac{1}{\tau}}$ in the integral.

$$\frac{1}{Dk^2+\frac{1}{\tau}} = \frac{\tau}{D\tau k^2+1} = \tau(1-D\tau k^2+D^2\tau^2 k^4-D^3\tau^3 k^6+...)$$

Putting this in Equation A.17 gives

$$< \rho(x,t)\rho(0,0) >_1 = \frac{a\tau}{4\pi}e^{-t/\tau}e^{-\frac{x^2}{4Dt}}\int_{-\infty}^{\infty}dk(1-D\tau k^2+D^2\tau^2 k^4-D^3\tau^3 k^6+...)e^{-Dt\left(k-\frac{ix}{2Dt}\right)^2}$$

$$(A.18)$$

The first term in the series of Equation A.18 is

$$\frac{a\tau}{4\pi}e^{-t/\tau}e^{-\frac{x^2}{4Dt}}\int_{-\infty}^{\infty}dke^{-Dt\left(k-\frac{ix}{2Dt}\right)^2} = \frac{a\tau}{4}\sqrt{\frac{1}{\pi Dt}}e^{-t/\tau}e^{-\frac{x^2}{4Dt}}$$

The second term in the series of Equation A.18 is

$$-\frac{a\tau D}{4\pi}e^{-t/\tau}e^{-\frac{x^2}{4Dt}}\tau\int_{-\infty}^{\infty}dkk^2 e^{-Dt\left(k-\frac{ix}{2Dt}\right)^2} = \frac{a\tau}{4}\sqrt{\frac{1}{\pi Dt}}e^{-t/\tau}e^{-\frac{x^2}{4Dt}}\left(\frac{\tau x^2}{4Dt^2}-\frac{1}{2}\right)$$

The third term in the series of Equation A.18 is

$$\frac{a\tau}{4\pi}e^{-t/\tau}e^{-\frac{x^2}{4Dt}}D^2\tau^2\int_{-\infty}^{\infty}dkk^4 e^{-Dt\left(k-\frac{ix}{2Dt}\right)^2} = \frac{a\tau}{4}\sqrt{\frac{1}{\pi Dt}}e^{-t/\tau}e^{-\frac{x^2}{4Dt}}\left[\frac{\tau^2}{4t^2}\left(\frac{x^4}{4D^2t^2}-\frac{3x^2}{Dt}+3\right)\right]$$

Adding the first, second, and third terms of the series together we have

$$< \rho(x,t)\rho(0,0) >_1 \approx \frac{a\tau}{4}\sqrt{\frac{1}{\pi Dt}}e^{-t/\tau}e^{-\frac{x^2}{4Dt}}\left[1+\frac{\tau x^2}{4Dt^2}-\frac{1}{2}+\frac{\tau^2}{4t^2}\left(\frac{x^4}{4D^2t^2}-\frac{3x^2}{Dt}+3\right)\right]$$

With both terms

$$< \rho(x,t)\rho(0,0) >\approx \frac{a\tau}{4}\sqrt{\frac{1}{\pi Dt}}e^{-t/\tau}e^{-\frac{x^2}{4Dt}}\left[1+\frac{\tau x^2}{4Dt^2}-\frac{1}{2}+\frac{\tau^2}{4t^2}\left(\frac{x^4}{4D^2t^2}-\frac{3x^2}{Dt}+3\right)\right]+G\tau^2$$

A.1.1   Numerical

The numerical model will be used to test the spatial correlation function, Equation A.14:

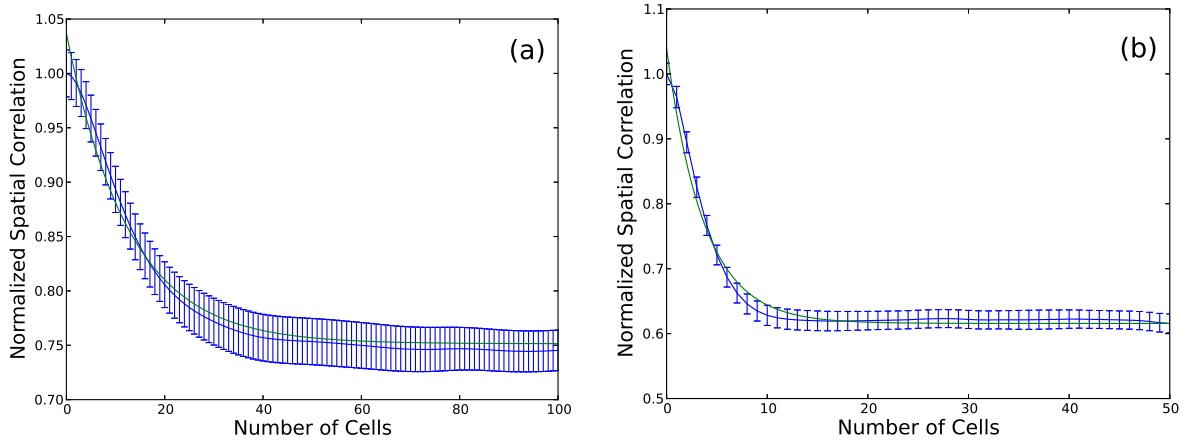$$< \rho(x,0)\rho(0,0) >= \frac{a}{4}\sqrt{\frac{\tau}{D}}e^{-x/\sqrt{D\tau}}+G\tau^2$$

Figure A.1: Normalized spatial correlation functions $< \rho(0)\rho(x) >$ in blue, with x measured as the number of cells. Normalization forced $< \rho(0)\rho(0) >= 1$. Green indicates fitted curve. (a) is for $D_M = 0.029/\pi$ s^{-1}, and found a fit to a general exponential $a + b \cdot e^{-(x-c)/d}$ with $a = 0.751392$, $b = 0.285534$, $c = 0$, and $d = 12.6625$. (b) is for $D_M = 0.29/\pi$ s^{-1}, and found a fit to a general exponential $a + b \cdot e^{-(x-c)/d}$ with $a = 0.615668$, $b = 0.422847$, $c = 0$, and $d = 3.70343$. Parameters are the same as those described in Section 3.2

In the numerical model (see Equation 4.1), two D_M values will be used: $D_M = 0.029/\pi$ s^{-1} and $D_M = 0.29/\pi$ s^{-1}. These are loss terms, and correspond to $1/\tau$ in the analytical model (see Equation A.1. $D_M = 0.029/\pi$ s^{-1} corresponds to a value of $\tau = 12.5$ cells and $D_M = 0.29/\pi$ s^{-1} to a value of $\tau = 3.94$ cells.

The numerical correlation functions for both of these loss parameter values have been plotted in Figure A.1. Also plotted in the figure are Marquardt-Levenberg fits to a general exponential function: $a + b \cdot e^{-(x-c)/d}$. The parameter d corresponds to the value of τ. For $D_M = 0.029/\pi$ s^{-1}, the fitted value of $d = 12.6625$ cells was quite close to the expected value of 12.5 cells. For $D_M = 0.29/\pi$ s^{-1}, the fitted value of $d = 3.70343$ cells was also close to the expected value of 3.94 cells.

A.2 Discussion

The three main results of the calculation are Equation A.14, for spatial correlations, Equation A.16, for temporal correlations, and Equation A.1, for spacetime correlations. The magnitude of all three of these of these functions is determined by the τ, which sets the magnitude of fixed nitrogen leakage.

The leakage sets the scale for the spatial correlations. In Equation A.14 the correlations go as $e^{-x/\sqrt{D\tau}}$. The numerical results shown in Figure A.1 show that correlations of this form and this dependence on τ are a good approximation. This accuracy tells us that the choice of white noise in Equation A.3, although artificial, is good enough for accurate spatial correlations. It also tells us that the absence of growth/advection in the model of Equation A.1 does little to affect spatial correlations. If the spatial correlations could be measured experimentally then, if the diffusivity was known, experiments could be used to measure the leakage in the cyanobacterial filament.

Furthermore, the leakage sets the scale for temporal correlations. In Equation A.16 the correlations go as $erfc(t/\tau)$. We have not measured the numerical correlations because they are not expected to be similar to the calculated correlations due to strong temporal correlations set by the devision time (see Section 6.4.2) that are not included in this preliminary calculation. This is for the same reasons that the there are temporal correlations in the spacetime pattern of starving cells, discussed in Section 6.4.4. Cells maintain a growth rate over a cell cycle and import of fixed nitrogen depends on the size of the cell. Both of these lead to temporal correlations lasting about a cell cycle. This would be very different from the white noise of Equation A.3. If these two parts of the model are removed, it is possible that Equation A.16 would describe the temporal correlations.

Appendix B

Permissions

copyright permissions letter Aidan Brown
Aidan Brown
to:
permissions
26/06/2012 18:56
Show Details

Hello IOP Permissions,

My name is Aidan Brown and I am currently working on my MSc in Physics at Dalhousie University in Halifax, Nova Scotia, Canada. I am planning on submitting my thesis and defending my MSc in August this year.

I am first author on two papers that have been published in Physical Biology this year:

Brown A I and Rutenberg A D 2012 Reconciling cyanobacterial fixed-nitrogen distributions and transport experiments with quantitative modelling Phys. Biol. 9 016007
Brown A I and Rutenberg A D 2012 Heterocyst placement strategies to maximize the growth of cyanobacterial filaments Phys. Biol. 9 046002

I am planning to use the work in these two papers as a large part of my thesis and require copyright permission in the form of a letter for the university to allow me to do so. If I have emailed the right place, could you please provide me with a copyright permissions letter? If not, could you direct me to where I could ask for one?

Thank you,
Aidan

PERMISSION TO REPRODUCE AS REQUESTED IS GIVEN PROVIDED THAT:

(a) the consent of the author(s) is obtained
(b) the source of the material including author, title of article, title of journal, volume number, issue number (if relevant), page range (or first page if this is the only information available), date and publisher is acknowledged.
(c) for material being published electronically, a link back to the original article should be provided (via DOI).

IOP Publishing Ltd
Temple Circus
Temple Way
BRISTOL   28/6/2012
BS1 6BE   Date   Rights & Permissions