THE UNIVERSITY *of* EDINBURGH

Modeling adaptive dynamics in microbial populations

with applications to the evolution of cellular resource allocation trade-offs

Christos Josephides

A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy
to the
University of Edinburgh

August 2015

Declaration

I declare that this thesis was composed by myself and that the work contained therein is my own, except where explicitly stated otherwise. The work has not been submitted for any other degree or professional qualification.

Christos Josephides, August 2015

Abstract

Adaptive evolution is the process by which natural selection, acting on variation within a population, promotes the survival of individuals that are more successful at reproducing and contributing to future generations. Evolutionary processes in microbes occur at the intersection of population genetics, natural selection, and underlying mechanistic constraints, to give rise to the repertoire of adaptation observed in nature. Understanding microbial adaptive evolution is of critical importance for human health for example, through the emergence of pathogenicity and antibiotic resistance. Moreover, the stability and function of natural and artificial ecosystems is contingent on the evolving interactions between microbes, and between microbes and the environment.

We present a modelling framework, based on the theory of adaptive dynamics, to investigate how cellular resource allocation trade-offs affect the adaptation process. We used resource-consumer theory, which explicitly models the interactions between cells and their environment, together with matrix models of structured populations, to implement phenotype-determined cellular strategies of resource allocation between mutually exclusive processes. We then analyse the outcome of competitions between different phenotypes across environmental and competitive conditions.

We applied our methods to the evolution of strategies (phenotypes) for resource allocation between two competing cellular process in microbial populations growing in chemostat-like environments. We calculated the adaptively stable strategies for several models and showed how state-structured population models can be mapped to simpler chemostat models on invariant manifolds. We then extended our analysis to the case where a limiting nutrient may be utilized using two alternative metabolic pathways. We described how the total population fitness of a metabolic strategy can be constructed from the individual decisions

of its constituent members. We developed numerical methods to simulate and analyse general models of adaptive dynamics using principles from graph theory and discrete Markov processes. The methods were used to explore the evolution of nutrient use strategies for microbial populations growing on two and three substitutable nutrients. We highlight the importance of the ancestral phenotype in channelling the adaptation process, which, together with the choice of the mutational kernel, influences the adaptively stable strategies and modes of co-existence. In a related finding, we show how some phenotypes are adaptively stable only in the presence of a competitor lineage that modifies the environment in a manner that permits another phenotype to invade. Our methods also reveal instances where historical contingency and chance have an important effect on determining the stable nutrient use strategies. Finally, we demonstrate the existence of adaptively stable periodic solutions whereby, under some conditions, phenotype successions are cyclical.

Our work builds on the foundation of adaptive dynamics theory to provide a general framework for analysing models of microbial adaptation. We focused on understanding the implications of underlying constraints and cellular resource allocation trade-offs in the context of adaptive evolution.

Lay summary

We constantly make decisions about how to spend our time, energy, and money. Like humans, cells have strategies for finding and using resources. For example, single-celled organisms, like bacteria, must choose the types of nutrients that they will consume. Decisions like these are critical for the survival of cells and organisms.

It is important to understand the survival strategies of single-celled organisms. For example, the survival of bacteria that cause disease is important for human health. Moreover, the survival strategies of all cells define how they interact with fragile and changing ecosystems.

Cells use biochemical means to sense their surroundings and implement their decisions. This biochemistry is encoded in their DNA, which is passed from a mother cell to its descendants. The DNA, therefore, can be thought of as implementing a strategy that directs the cell's actions.

Evolution occurs through errors in the transmission of DNA from the mother cell to its daughter cells; as a result, daughter cells may have an altered decision-making strategy. In most cases, the new strategy will not be as effective as that of the mother cell. The daughter cell that uses the altered strategy, then, will not be as capable at seeking resources and using them in an efficient manner. Rarely, however, the new strategy will perform better than the previous one. The daughter cell that uses the better strategy will be able to find more resources and ultimately produce more offspring. In this manner, then, the numbers of cells using the old strategy will begin to decline as the number of cells using the new strategy will gradually become more abundant.

We investigated how bacteria that use different survival strategies interact with each other. We used mathematical models and computer simulations to create

digital cells. We then grew the cells in a variety of simulated natural and laboratory conditions. The strategies were encoded in a digital DNA that would, occasionally, have a mutation (copy error). This error would give rise to a new strategy in the cell that would alter its ability to survive.

We found that the degree to which a strategy increased cell survival depended on two factors: first, on the environmental conditions that the cells grew in; and, second, on the type of competitor strategies that were grown together with this strategy. Using our models, we were able to predict when a new strategy would out-compete an older strategy and drive it to extinction. We also found that, under appropriate environmental conditions, strategies that specialize in using resources in unique ways could co-exist. This co-existence is one way in which diversity is maintained in ecosystems.

In conclusion, our work uses mathematical and computer methods to model evolution in single-celled organisms. We find ways to describe and predict how the survival strategies of cells will change over time, in response to competition and environmental change.

Acknowledgments

I would like to extend my gratitude to my supervisor, Prof. Peter Swain, for his unwavering support throughout my PhD. For our enthusiastic - and sometimes endless - discussions, and for his general mentorship that contributed so much to my intellectual and personal growth these past few years.

Many thanks are also owed to current and past members of the Swain lab. Bruno, Catie, and Ivan warmly welcomed me to the group, and to Edinburgh, four years ago and made me feel at home in spite of my naivette. Elco and Matt became good friends with whom I've shared the ebb and flow of professional and personal life.

I've learned a lot from Andrea Weisse during her post-doc at Peter's lab and I give my thanks for her motivation, insightful recommendations, and particularly her willingness to listen to and discuss my own ideas. Nikola Popovic helped us tremendously with understanding the invariant manifolds and dealing with other troublesome facets of the work, and for that I am deeply grateful.

I could not have attained any measure of achievement without the love and support of my parents, George and Elizabeth, and of my wonderful sisters, Malvina and Markella, who have encouraged me at every stage of my life.

Finally, it would have been a very different journey without Alisa, who has been by my side during the best of times and the worst of times. Whose kindness and patience have been a place of solace, and her intelligence and curiosity a source of boundless inspiration.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction to the introduction

Adaptive evolution is the process by which natural selection, acting on variation within a population, promotes the survival of individuals that are more successful at reproducing and contributing to future generations. Theoretical and experimental work has revealed how evolutionary processes take place at the intersection of population genetics, natural selection, and underlying mechanistic constraints, giving rise to the repertoire of dynamics observed in nature and in controlled laboratory evolution experiments. Recent efforts have accumulated evidence concerning the potential of cellular and organismal trade-offs to shape evolutionary outcomes. The universality of fundamental trade-offs, arising from finite resource allocation between competing and mutually exclusive processes in the cell, has been demonstrated at various organisational levels and the consequences for adaptive dynamics in human and non-human systems are actively being explored.

Evolutionary processes involve feedbacks from multiple timescales and emerge from non-linear interactions between organisms, and between organisms and their environment. Because of these complications, theoretical treatments of adaptive evolution usually focus on specific aspects of adaptation while making restrictive assumptions about components that are not explicitly addressed. A conceptual and modeling approach that integrates ecological interactions, underlying mechanistic constraints, and adaptive dynamics inspired by evolutionary game theory,

will be more successful in reproducing and predicting the scope and trajectories of phenotypic adaptation relative to the more restrictive methods. The need to rigorously and comprehensively investigate adaptive evolution is emphasized by its central role in the emergence of pathogenicity, in the development of antibiotic and antiretroviral drugs, and in cancer progression, since these processes involve the adaptation of clonally-replicating populations of cells to shifting selection pressures. Moreover, with the advent of synthetic biology, there is a pressing need to understand the limitations, constraints, and selective pressures on deployed synthetic circuits that are contingent on host organisms and environmental circumstances - both of which can potentially drive function toward unintended directions.

To partly address the need for a more comprehensive theoretical treatment of evolutionary processes we propose a modeling framework based on the theoretical underpinnings of adaptive dynamics theory. Our framework integrates ecological realism, by making explicit the interactions between organisms and between organisms and their environment; phenotypic limitations, derived from underlying mechanistic constraints; and competition/invasion analysis inspired by evolutionary game theory. The approach is applied to the adaptation of cellular resource allocation between two resource-intensive operations in the cell, to the evolution of metabolic pathway specialization, and to evolution of nutrient selection. We present analytical results that extend the classical theory and, for more complicated models where analysis is intractable, we develop novel numerical procedures to streamline and generalize the analysis of adaptive evolution using principles from graph theory and discrete Markov processes.

1.2 Adaptive evolution

1.2.1 What is adaptive evolution and why is it important?

Adaptive evolution is the process by which natural selection, acting on genetic variation within a population, promotes the survival of individuals that are more efficient at reproducing; in other words, they are better adapted to their environment. While this is a naive definition of adaptive evolution, it nevertheless raises two fundamental questions [1] about this process:

First, what is the genetic basis of variation? Second, how does selection favor the more well-adapted genotypes from among this variation?

The study of the processes that generate variation has a long history whose origins can be traced back to the era before molecular biology. Perhaps most recognizable is the work of Mendel in the 19th century on the laws governing the inheritance of phenotypic traits. The discovery of the genetic code and the recognition of DNA as the molecular form of genetic information, together with a theoretical treatment of selection (notably by Fisher), led to a search for molecular mechanisms by which genetic variation arises within populations.

The development of the neutral theory of molecular evolution, originally proposed by Kimura in the 1960s [2, 3, 4], set the stage for subsequent research into the molecular forces that drive genetic variation. The most salient element of the neutral theory is that, at least at the molecular level, most observed random DNA mutations are neutral - in the sense that they do not affect an individual's survival and reproduction efficiency.

The emergence of a beneficial mutant allele is not sufficient to guarantee a change in the population's genetic variation in the long term [5]. Individuals bearing the mutation must first escape extinction through a largely stochastic birth-death process - necessarily, since the allele is initially very rare in the population. The probability of fixation depends on the size of the population as well as on the selective advantage of the mutation.

The Moran process [6] is a stochastic process that can be used to model the fate of a mutant allele in a population of finite and fixed size. Suppose that the selective advantage of a rare mutant allele relative to an abundant resident allele

is r, defined as the ratio of the per-capita growth rate of the mutant compared to that of the resident. In the absence of frequency-dependence effects, it can be shown that the probability of eventual fixation is

$$q_i = \frac{1 - r^{-1}}{1 - r^{-N}}$$

Where i is the number of mutant alleles at the start of the process. If the mutation is neutral, i.e. $r = 1$, we have

$$\lim_{r \to 1} q_1 = \frac{1}{N}$$

This means that the probability that a neutral mutation fixes, starting with a single mutant copy, is inversely proportional to the population size. For non-neutral selective advantages, the probability of fixation depends non-linearly on the population size and the mutant's initial frequency. An allele's selective advantage [7] is a measure of how strongly natural selection can favor the replication of this allele relative to other alleles, or relative to the population's mean; therefore, strongly beneficial mutations have a higher probability of fixation, or of being maintained, in the population. In fact, the distribution of selection coefficients in a population's supply of beneficial mutations is currently actively explored with various theoretical and experimental models [8, 9].

While the foundations of the molecular mechanisms underlying the various random mutational processes have since largely been elucidated [10], driven by the recognition, and contention, of the neutral theory, the inescapable implication that phenotypic evolution (unlike molecular evolution) must necessarily involve the interaction of population genetics, natural selection, and biophysical constraints, has created a rich field of scientific research [11, 12, 13]. An illustrative and important example is provided in the case where the selective advantage of a phenotype depends on its frequency; for example, it has been demonstrated theoretically [14,15] and experimentally [16,17,18] that negative frequency-dependent selection can be a mechanism by which multiple phenotypes can be maintained in the environment because the selective advantage of a particular allele is negatively correlated with its frequency.

The emergence of a beneficial mutant allele in the population is not sufficient to guarantee increased genetic variation in the long term [5]. Individuals bearing

the mutation must first escape extinction through a largely stochastic birth-death process - necessarily, since the allele is initially very rare in the population. The probability of fixation depends on the size of the population - more correctly, on the effective population size. Once the allele has escaped accidental extinction, defined (somewhat) arbitrarily as reaching some number of individuals, below which the frequency of the allele can only be treated as a random walk, the individuals bearing the allele will begin to more directly and deterministically compete with other individuals in the population. The allele's selection coefficient [7] is a measure of how strongly natural selection can favor the replication of this allele relative to other alleles, or relative to the population's mean; therefore, strongly beneficial mutations have a higher probability of fixation, or of being maintained, in the population. In fact, the distribution of selection coefficients in a population's supply of beneficial mutations is currently actively explored with various theoretical and experimental models [8, 9, 9]. Finally, the selective advantage of the allele usually depends on its frequency; for example, it has been demonstrated theoretically [14, 15] and experimentally [16, 17, 18] that negative frequency-dependent selection can be a mechanism by which multiple phenotypes can be maintained in the environment because the selective advantage of a particular allele is negatively correlated with its frequency.

Adaptive evolution is central to many areas of biology for the insight it provides into processes that shape human biology and non-human systems that affect human biology. In particular, a recent focus of biomedical research has been to understand the adaptation processes involved in increased pathogenicity (for example, in the HIV epidemic [17]), drug resistance (particularly antibiotic [19], antiretroviral [20], and chemotherapy [21] drugs), and in cancer progression [22]. This particular area deals with adaptive evolution in the narrower scope of clonally reproducing cells - that is, the entirety of an offspring's genetic material is derived solely from a single parent.

A rigorous treatment of adaptive evolution entails more than the enumeration of the mutations that have emerged and established. More carefully, mutation identification must be combined with quantitative modeling and tracking of lineage composition to determine the fate and trajectory of mutations, phenotyping to establish the selectivity of mutations, and characterization of phenotype-phenotype and phenotype-environment interactions to understand and quantify the parallelism and determinism of the observed lineage dynamics.

1.2.2 Experimental investigations of adaptive evolution

Experiments related to the adaptive evolution of clonally-reproducing organisms have been carried out since the pioneering work by Luria and Dellbruck [23], which demonstrated that the emergence of phage-resistant bacteria was not a result of mutation induced by the co-cultured phage virus but rather through selection on existing genetic variation within the population. This and other important early work(notably [24], and reviewed in [25]) already contained the elements and tools that are still used today to study adaptive evolution; in particular, a combination of controlled laboratory environments that microorganisms are propagated in, phenotype characterization, analysis with dynamical population models, statistical modeling of underlying mutation processes, as well as an appreciation of the stochastic nature of the interacting forces at play.

Experiments for observing adaptive evolution are conceptually simple. The experiment is initialized with one well-characterized lineage to establish a common genetic starting point. The cells are grown in a defined and controlled environment, usually with some kind of stress like limitation for a necessary resource, and are propagated either via continuous growth in chemostat vessels or via batch-growth by serial transfer in growth flasks. Samples are taken periodically and are assayed for phenotype characterization and are now also usually sequenced to some extent; in addition, samples are kept indefinitely using cryopreservation. Revived samples can serve as reference points from which the experiment may be restarted, and can be used in comparative studies against the ancestor or other samples taken from the past or future of the laboratory evolution experiment.

The particulars of the experiment depend on the objective of the investigation. Mutation accumulation studies, for example, seek to measure the rate and distribution of mutations and their effects [26, 27]. These studies typically involve propagating the population through severe bottlenecks to eliminate genetic variation that selection can act on, thereby allowing most or all mutations to persist and be counted. Most studies on adaptive evolution, however, strive to preserve some, if not all, of the genetic variation in order to investigate the selective forces acting on it.

The simplest of adaptive evolution experiments can be thought of in terms of the iterated emergence of mutations and the resolution of their eventual extinction or fixation (dominance). This pattern of selection was originally reported in 1951 by

Atwood et al. [28] after the observation that, periodically, a rare mutant would emerge that would sweep through the population until it replaced all the pre-existing lineages. However, it was quickly determined - indeed, even suspected in the paper by Atwood et al. itself - that periodic selective sweeps represent a very small fraction of the repertoire of adaptive dynamics. Extensive work has since demonstrated the importance of clonal interference, frequency-dependent selection, epistasis, and other factors that determine the degree of parallelism and repeatability in evolution (discussed below).

Several recent studies, employing high-throughput "barcoding" and sequencing methods, have tracked thousands of co-evolving lineages - from emergence to extinction or fixation - in a single experiment. These results suggest that the contribution of clonal interference (defined as the simultaneous competition between multiple beneficial mutations that are present in distinct lineages) in adaptation must be even greater than previously thought [29,30]. The effects of clonal interference in clonally-reproducing organisms are a subject of ongoing research [31]. As a general rule, clonal interference is expected to decrease the rate of adaptation, since there is always one mutation that is the most beneficial from among the pool of transiently coexisting beneficial mutations; moreover, it is expected that the mutations of highest benefit would sweep first, followed by mutations with gradually diminishing benefits. These effects have been predicted theoretically and verified experimentally.

The effects of frequency-dependent selection in the metabolic diversification of a population of evolving bacteria, growing in a medium containing one optimal (glucose) and one sub-optimal (acetate) carbon source, was recently investigated [32]. In this work a clonally homogeneous population favoring growth on glucose, but also able to utilize acetate following glucose depletion, diverged to two lineages after a few hundreds of generations: a glucose "specialist" and an acetate "specialist". It was concluded that frequency-dependent competition for the carbon sources, that were depleted at different rates, resulted in niche construction, which in turn created the conditions that led to divergence of a "generalist" ancestor to two "specialist" lineages; moreover, the experimental results were in agreement with theoretical predictions from adaptive dynamics theory. In a follow-up investigation, the selective consequence of mutations was further examined - in the context of genetic epistasis and frequency-dependent selection - during a laboratory adaptation experiment [33]. By revisiting samples taken during the

metabolic diversification experiment, and swapping ancestral with mutant alleles in genomic backgrounds derived at different time points of the experiment, it was shown that the selective effect of a mutation is highly contingent on the genomic context as well as on the competitive environment.

Experimental evolution of microbial resistance to a particular antibiotic, known to be mediated by five point mutations in a single gene, also highlighted the importance of epistasis (whereby the effect, or benefit, of one mutation is contingent on the presence of other alleles) on the selective benefit of mutations [19]. Through a systematic construction of the $2^5 = 32$ possible alleles, and using a characterization of the resulting phenotype (antibiotic resistance) as a measure of the selection coefficient, it was shown that the majority of the mutational trajectories (102 of $5! = 120$) leading from the wild-type allele to the resistant allele are effectively inaccessible because extensive epistasis precludes the succession of a lineage by a descendant lineage with a lower selection coefficient. Other notable studies employing a similar approach based on systematic reconstruction and phenotyping are reviewed in reference [34] and represent an emerging way of thinking about, and experimentally investigating, empirical fitness landscapes; it should be noted, however, that a limitation of these studies is their persistence (some of it necessitated) in adopting static fitness landscapes when, in reality, frequency-dependent selection implies that the fitness landscapes are more likely to be dynamic and dependent on the pre-existing competitors.

The repeatability, or parallelism, of the observed outcomes during laboratory evolution experiments is another aspect of the adaptive process that has received much attention. A highly restricted mutational landscape, such as the one reported for the evolution of antibiotic resistance, suggests that adaptation might have a largely deterministic character [19]. The outcome of the metabolic diversification experiment was reproducible across multiple replicates [32]; in fact, in a subsequent study that more closely examined the mutational trajectories of the diversifying lineages it was found that similar patterns of selective sweeps took place across experiment replicates [35]. This was taken as strong evidence that not only are the specific mutations highly parallel but that the entire process of adaptation might be strongly deterministic - including the pattern of frequency-dependent selection, the shaping of the fitness landscapes, and the ecological and demographic dynamics. Nevertheless, some studies have pointed out that the effect of historical contingency and chance during the adaptation process can be

pronounced, particularly during the later stages of the adaptation process [29]; consequently, the repeatability and parallelism of adaptation remains an open and actively pursued question.

1.2.3 Theoretical frameworks for modeling adaptive evolution

The quantitative analysis of evolutionary processes has a long history and many disparate theoretical treatments of ecological, biological, and genetic processes have ultimately been synthesized to comprise the contemporary analytical tools that exist today.

Ecological modeling of natural populations, as recognized today, probably began with the formulation of prey-predator dynamics by Lotka and Volterra in the early part of the 20th century [36]. The objective then was to determine the trajectories of the numbers of two interdependent species - a predator and a prey - through time and subjected to certain parametrizable characteristics. The Lotka-Volterra equations are themselves the conceptual successors of a class of models employing the logistic function, which was first studied by Verhulst in relation to the growth of self-limiting biological populations, as originally conceptualized by Thomas Malthus [37].

Modeling the dynamics of microbial populations in well-mixed environments became a research topic of much interest with the development of the chemostat by Novick and Szilard in 1950 [24, 38], which was itself motivated by the earlier experiments on microbial growth under nutrient limitation by Monod [39]. The chemostat made possible the uninterrupted culture of microorganisms under controlled and reproducible conditions with continuous inputs and outputs, and has since been used as the laboratory analog for natural, open, systems [40]. Perhaps more importantly, at least from a theoretical perspective, it motivated extensions to the original "one-population/one-nutrient" model and as a result led to the formulation of a mathematical theory of nutrient competition with many important industrial and biomedical applications [40, 41].

The chemostat model became the foundation for many variations that subsequently posed, and sometimes answered, more realistic biological questions and interesting mathematical problems. For example, a classic result from a standard

dynamical system analysis of chemostat models is the competitive exclusion of all lineages by the lineage with the fastest growth rate [42]. This surprising result, which holds for a well-mixed chemostat and a single limiting nutrient, is at odds with the organismal diversity and clonal adaptive radiation observed in the wild and in chemostats [43], and was originally reported as the "paradox of the plankton" for its contradiction with the great diversity of aquatic plankton feeding on a single resource [44]. The proof of the competitive exclusion principle [40, 45] was followed by an outpouring of experimental work that both sought to confirm (collected in [46], notably [47]) and invalidate the theory [48, 49, 50].

In response to experimental rejection of the universality of the competitive exclusion principle, new theoretical approaches emerged to explain the observed diversity of organisms. Levin demonstrated how spatial and temporal heterogeneity in the environment can generate biodiversity [51, 52] while May and MacArthur argued that the overlap of niches and competition in the real world is never perfect [53]. Others have argued using non-equilibrium arguments, stating that, for closely-matched organisms, the rate of exclusion will be so slow that it will approach speciation time [54, 55], while others have argued for more nuanced modes of competition, emphasizing intraspecific competition [44, 56, 57], and self-renewing (i.e. biotic) resources [58].

The advent of improved molecular technologies, particularly DNA sequencing, saw the application of the theoretical tools that had been developed for the analysis of evolutionary dynamics at a coarse level to the modeling of populations and individuals characterized by specific nucleotide sequences, and promoted the recognition of mutation as an important component of adaptation models. To be use a more concrete example, the extensive genetic heterogeneity exhibited by viruses and some bacteria led to the concept of a "quasispecies", originally by Eigen and Schuster in the 1970s [59]. A quasispecies is defined as a distribution of closely-related mutants that are generated and maintained by mutation-selection balance [60]. Quasispecies models have dependent variables for the frequency of genetically homogeneous individuals, much like chemostat models, but also include a mutational processes that represents the erroneous replication process that allows a small fraction of the reproducing population of one mutant to produce offspring in another mutant population. The system of differential equations can be solved to yield a stationary distribution of mutants using standard linear algebra methods. Because the rate of mutation is not very small rela-

tive to the rate of reproduction, and because there are only a few "genotypes" that are accessible from each other, the peak of the distribution - which represents the fastest-growing mutant - can support a cloud of mutants that have lower replication rates. This method for generating and maintaining genotypic diversity was successfully applied to mechanistic pathogenicity models in human immunodeficiency viruses [60, 61, 62], and more recently extended to account for antiretroviral-induced viral dynamics [20] within the host as a response to the emergence of resistance to antiretroviral drugs within a host.

Quasispecies modeling has also been successfully used, in combination with the replicator equation [63], to investigate the generation and maintenance of bacterial diversity in a chemostat containing a single nutrient [64]. The model allowed for variation in the metabolic mode of nutrient use, in the form of trade-offs (for example, metabolic rate decreased with increasing metabolic yield), and postulated the existence of discrete lineages grouped according to phenotype corresponding with their metabolic mode. It was shown that, for sufficiently high mutation rates, that the co-existence of two quasispecies is indeed possible provided that one quasispecies replicates at a fast rate, but is associated with a "narrow" distribution of mutants, while the other replicates more slowly but has a "flatter" distribution - indeed, this result was aptly named the "maintenance of the fittest and the flattest". The coexistence was attributed to quasispecies-level frequency-dependent selection.

It is important to distinguish the quasispecies-type models, where the dependent variables are the frequencies of a finite set of fixed and pre-determined (genetic) types, from another class of models where the dynamical variables are themselves the parametrized phenotype trait values. The phase space for the latter class of models is the space of all possible phenotypes; in contrast, the former class of models never goes beyond the set of types initially provided by the model. However, it is also important to acknowledge that, as will be discussed next, both types of models present a paradigm where the classic static fitness landscapes have been replaced by dynamic fitness landscapes that shift as a function of the composition of phenotypes in the environment [36]. This is reminiscent of - and, indeed, inspired by - the frequency-dependent fitness of evolutionary strategies in evolutionary game theory proposed by Maynard Smith in the 1970s [65, 66, 67] whereby the "payoff" of a particular behavior or trait is dependent on the probability of encountering each of the other behaviors and traits in the environment.

The framework of adaptive dynamics, originally formulated by Hanz Metz [68, 69, 70, 71] in the 1990s, and expanded by many others [72, 73, 74, 75], provided an alternative to quasispecies modeling that succeeded in, amongst other achievements, demonstrating the theoretical possibility for diversification without geographical isolation and without requiring the high mutation rates necessitated by the mutation-selection balance of quasispecies modeling. The conceptual underpinnings of the framework are discussed below.

This type of analysis begins with the usual formulation of ecological models that describe the "demography" of a population consisting of clonally-reproducing individuals, which are treated as the *resident* lineage in the environment, much like in traditional chemostat models. Importantly, this lineage is characterized by a phenotypic trait, or traits, that is somehow parametrized and all individuals of the resident lineage are assumed to possess the same trait. The analysis proceeds by showing the existence, uniqueness, and global attractiveness of the dynamical system's steady state.

The fundamental unit process of adaptation in this framework is the emergence of an initially scarce mutant lineage, characterized by a novel phenotypic trait parameter. The fate of this nascent mutant lineage can be determined by its initial growth rate; importantly, this initial growth rate is conditioned on an environment modified by the resident lineage(s), as well as by the presence of these lineages. The initial growth rate, sometimes known as the invasion exponent or invasion fitness, can then be written as a function of the mutant and resident phenotype parameters. The sign of the invasion exponent predicts the long-term fate of the mutant lineage: if the exponent is positive then the mutant lineage will establish; otherwise, if the sign is negative, the lineage will eventually become extinct. It should be emphasized that the invasion fitness function is derived from the particulars of the ecological model - contrary to most approaches in evolutionary dynamics where fitness functions are phenomenologically obtained or assumed.

The derivative of the invasion fitness function, or of a linearized form of the function, can be taken with respect to the phenotype parameter. Because the theory of adaptive dynamics only deals with gradual phenotypic change, the derivative is evaluated for a mutant trait approximately equal to the resident trait.In practice, this is done by linearizing the mutant's invasion eigenvalue expression, which is a function of the resident and mutant phenotype parameters, and evaluating the

first-order term (derivative) for the case where the mutant phenotype is equal to the resident phenotype. The result is a selection coefficient in the form of an ordinary differential equation in the phenotype parameters over time; time, in this case, is interpreted on an evolutionary timescale. The resulting dynamical equation gives the trajectory, in phenotype space, of the dominant trait in the population as mutations emerge and sweep to fixation.

In effect, the adaptive dynamics approach combines population dynamics, which occur at a demographic, or ecological, timescale, with evolutionary dynamics while emphasizing the effect of frequency dependent selection that is a hallmark of evolutionary game theory. The appeal of the adaptive dynamics framework is that it can be applied independently of the model structure; in addition, it is fairly straightforward to extend the basic workings of the framework.

On the other hand, adaptive dynamics has been criticised for some of its assumptions [76, 77, 78, 79, 80], notably: how valid is the assumption of a continuum of phenotypic effects; how does the theory respond to mutations where the difference between the mutant phenotype and the resident phenotype is not small, and how can more realistic mutational models fit with the theory; finally, the assumption that the emergence of a mutant lineage is well-separated in time from the emergence of another mutant lineage is suspect given the prevalence of clonal interference observed in many experimental studies. Several attempts have been made to reconcile these criticisms [81, 82, 83, 84] as the framework matures and the more troublesome aspects of the theory are clarified and improved. In this work we will develop a modeling framework based on the foundations of classical adaptive dynamics theory that departs from the restriction of continuous phenotype spaces and facilitates the characterization of the adaptation process under various adaptation/mutation kernels.

1.3 The economy of the cell: resource allocation and tradeoffs

Living organisms, conceptually abstracted as self-contained systems with regulated inputs and outputs, face a common and fundamental problem: given finite resources and a number of processes that compete for these resources, how should the resources be allocated? [85,86] The intuitive answer is that organisms should allocate resources in such a way that maximizes their reproductive efficiency; indeed, similar questions and answers have been pursued in the context of autocatalytic chemical reactions by Hinshelwood and Koch [87] since the 1950s. There are many parallels between this 'economic' terminology and evolutionary game theory that seeks optimal strategies in the context of evolving populations. In this section, I will discuss the concept of the economy of the cell, resource allocation and tradeoffs, and the implications of this emergent way of thinking about biology to adaptive dynamics and evolutionary theory.

One of the earliest and most appealing instances of economic thinking in biology can be found in the evolutionary theory of ageing proposed by Kirkwood in the 1970s [88,89]. In this, now seminal, work it was proposed that ageing arises due to the accumulation of errors in the synthesis of macromolecules that eventually leads to an "error catastrophe" manifested as the deterioration of function and increased risk of mortality associated with ageing. It was also known at the time that fidelity of macromolecule production relied on several quality control mechanisms that were costly to maintain [90] and that the efficiency of these pathways was not consistent between organisms, or even between somatic and germline cells in the same organism. Kirkwood proposed that the organismal and cellular investment into repair and maintenance pathways must come at the expense of other functions that contribute to reproductive fitness and provided several theoretical and observational examples. It was shown using population models that there must exist an optimal amount of repair investment; moreover, this optimum should be inversely proportional to the degree of extrinsic mortality (for example, through predation). One prediction of the theory is that animals whose evolutionary past entailed a relatively low risk of death from external factors should invest more in repair because such investment would prolong their reproductive capacity and increase their overall fitness; in contrast, animals that have evolved with a high risk of extrinsic mortality are not expected to invest heavily in so-

matic repair because this would not guarantee a significantly increased number of offspring (or life span). It is taken as a consequence of this theoretical prediction that "aged" animals are rarely encountered in the wild when these are the subject of predation, but that the same animals raised in protected environments age and have lifespans that correlate with survival rates in their natural environments [91].

The economic paradigm has at its root fundamental biochemical and physical tradeoffs that have a measurable impact on the fitness of the organism - and, as will be discussed later, are expected to be important in the evolution and adaptation of biological systems. Several experimental studies in recent years have adopted this paradigm, ranging from investigations into the optimality of expression of a single protein, to flux-balance analysis of entire metabolic pathways, and whole-proteome analyses resulting from mRNA competition for finite ribosome resources.

In an elegant experiment by Dekel and Alon [92] the cost, in terms of growth rate, of expressing the lac operon was measured in bacteria by artificially inducing gene expression in the absence of lactose; similarly, the expression benefit was measured at full activation as a function of external lactose. Laboratory evolution of bacteria at different levels of external lactose determined that bacteria adapt toward the theoretically determined cost-benefit optimum at a wide range of lactose concentrations. While the authors made no attempt to attribute the source of the cost to particular factors in this study, it has been calculated by others that gene expression (both transcription and translation) is energetically costly to the cell [93]. Energy in the cell can be measured in terms of high-energy phosphate bonds and an accounting of the phosphate bonds involved in the synthesis of macromolecules shows that an increase in gene expression has a selective disadvantage - assuming no concomitant benefit - that is high enough to be acted on by natural selection in typical sizes of bacterial populations.

Tradeoff and optimization approaches have been employed at the level of metabolic networks to demonstrate how optimal microbial growth rates on common carbon substrates can be predicted from in-silico models. Using the complete genome sequence and annotation of the bacterial gene the stoichiometric matrix of bacterial metabolism was constructed and constraint-based flux-balance analysis was used to predict the optimal growth rate of bacteria grown under several selection

pressure regimes [94, 95]. An important concept that took hold from this type of analysis is that it may be possible to define a phenotype subspace that is the set of all optimal behaviors.

The universality of phenotype trade-offs and multi-objective optimization was proposed in a recent paper by Shoval et al. [96]. This work relies on the assumption that organisms perform multiple tasks that contribute to fitness and that fitness must be an increasing function of performance in all tasks; a given phenotype, however, cannot perform equally well at all tasks. The Pareto front, a concept from economics and engineering, is the set of phenotypes that contains no points that can be be outperformed by any other point at all tasks, is a natural way to define and calculate these subspaces. The authors used data on the transcriptional activity of thousands promoters in bacteria as these transitioned from exponential to stationary phase, and showed that genes for growth (ribosomes, transcription, translation) were on a tradeoff line against genes for stress and survival and that expression transitioned from growth-optimal allocation of transcriptional resources, during exponential growth, to survival-optimal allocation during stationary phase.

In principle, a tradeoff between two or more tasks may arise whenever these tasks 'compete' for a finite resource. Elegant theoretical and experimental work by Matthew Scott and others [97] has demonstrated how ribosomal capacity can be thought of as a finite resource to be partitioned between classes of gene products - namely, ribosomal proteins, a group of growth-rate-invariant proteins, and a group of proteins whose expression depends on the growth rate. Using a simple phenomenological model it was shown that optimal ribosomal allocation depends on the nutrient quality: bacteria divert resources away from transcribing non-ribosomal proteins - including proteins that sequester nutrients from the environment and would otherwise provide growth-rate benefits - toward ribosomal proteins as a growth-rate maximizing strategy. The optimum allocation of ribosomal resources was measured as a function of nutrient quality and used to derive strikingly simple bacterial growth laws [98] with important predictive utility. More recently, the bacterial growth laws were re-derived using a minimal mechanistic cell model that explicitly incorporates tradeoffs arising from competition for energy, competition for finite ribosomal resources, and competition for a finite proteome size [86, 99]. Fundamental properties of microbial growth, such as gene dosage compensation and host effects on the performance of synthetic

circuits, could be investigated using a mechanistic model in the context of these cellular tradeoffs.

1.3.1 The impact of trade-offs and constraints on adaptive evolution

The potential for resource allocation and tradeoffs to shape evolutionary outcomes is becoming widely recognized and recent work in this field has accumulated evidence promising to improve and inform our understanding of adaptive evolution; moreover, it is also becoming accepted that diversity amongst extant organisms, and diversity within the same organism can be understood in terms of underlying multi-objective optimization processes. A conceptual and modeling framework to synthesize cell economics, particularly tradeoffs arising from resource allocation constraints, embedded in ecological models that explicitly define organismal and environmental interactions, and analyzed within the framework of adaptive dynamics theory, is, however, lacking.

Adaptive evolution involves the overlap of processes that operate at different timescales which interact in a highly non-linear way. A unifying framework that incorporates ecological realism, mechanistic constraints, and adaptive dynamics can aid in understanding how the ecology of an organism, together with its internal tradeoffs, shape the phenotypic evolution of characteristics and behaviors. Modeling adaptive evolution can reveal the repertoire of adaptive dynamics that are possible from which we can infer the range and diversity of extant phenotypes that are likely to be seen in nature and in the laboratory. Ecological realism in the models permits the investigation of how changes in the ecology of an organism modulate the process of adaptation, with important implications for pathogenicity, disease progression, and ecosystem creation and destruction. In addition, with emergent synthetic biology approaches, it may be possible to alter or skew the internal tradeoffs and resource allocation of a cell; for example, when overloading the cell's transcriptional machinery with a highly expressed gene product. Because synthetic circuits are not exempt from mutation and selection, an integrated modeling approach can grant the predictive power to estimate the rate and direction of evolutionary change expected to take place as a consequence of deploying a synthetic circuit within a particular host and environment.

A unified approach must also address the shortcomings and limitations of the existing methods, either through improvement of individual components or through complementarity by additional included components. Replicator-mutator models, which combine the quasi-species concept with the replicator equation, are

not always satisfactory because of the high mutation rates and/or weak selection required [100,101] to maintain diversity in the mutation-selection balance regime. In addition, such models have been criticized with regards to the limitation that phenotype traits are confined to a pre-set discrete range of values - although this may, in some cases, be warranted. The theory of adaptive dynamics, on the other hand, uses a continuous phenotype phase space, and does not require high mutation rates to predict the generation and maintenance of phenotypic diversity [102].

It is important to appreciate that the theory of adaptive dynamics models phenotypic evolution while being agnostic about the underlying mutational process and genotype-level change. Recent work in experimental evolution has demonstrated the transient co-existence of a large number of lineages during a laboratory evolution experiment [29]; however, there was considerable overlap between the fitness of many of the lineages which implies that change at the genotype level may not map onto a significant phenotypic change. Doebelli and others [32, 35, 103] use the "ecotype" terminology to interpret their adaptive evolution experiments and theory in a manner that may first appear contradictory to these experiments: an ecotype may be defined as a group of closely-related types that, within themselves, obey the principle of competitive exclusion, but different ecotypes need not compete to the exclusion of all ecotypes except one. In other words, two lineages with phenotypes that belong to the same ecotype have similar enough behaviors or niches and will compete more severely for resources and growth, whereas two lineages belonging to different ecotypes occupy significantly different niches and can stably co-exist. Unlike quasispecies theory, however, diversity is not maintained via mutation-selection balance, and even more importantly, mutations from one quasispecies to another cannot be responsible for the maintenance of a stable coexistence. Theoretical work by Wagner and others on neutral genetic networks [104] provide a framework within which extensive mutational diversification is prevalent. Such diversity is experimentally supported in 'barcoded' lineage tracking experiments, and does not necessarily manifest a concomitant phenotypic diversity. Additional work on phenotypic "capacitors" has revealed the presence of gene products that buffer genetic (and environmental) variation [105,106]. This permits the accumulation of genetic polymorphisms that nevertheless do not significantly affect phenotype, and therefore fitness; however, a failure or overloading of buffering capacity provides a mechanism by which phe-

notypic variation may express. Therefore, the phenotype (or ecotype) centered approach taken by the theory of adaptive dynamics does not preclude, nor is it invalidated by, extensive genetic polymorphism and transient co-existence of distinct lineages observed experimentally.

Models of population and adaptive dynamics frequently ignore ecological interactions between organisms and between organisms and the environment. As a result, many modeling efforts have relied on assumed fitness functions, whose form may be chosen largely as a matter of analytical convenience. Explicitly defining ecological interactions, on the other hand, allows fitness to emerge naturally from these interactions - for example, as a result of competition for nutrients; although the fitness function may still be parametrized, the parameters usually correspond to experimentally measurable ecological aspects, and, in the case of laboratory evolution, aspects that are under experimental control. Therefore, particularly for investigating how ecology affects adaptive evolution, a unified modeling framework must allow for more natural interpretations of the ecological interactions likely to determine organismal fitness.

Finally, a method that allows the inclusion of tradeoffs between cellular processes, but that does not severely hamper analytical tractability or require extensive mechanistic models is sought. The analysis of adaptive dynamics requires analytical manipulations, or numerical computation, that preclude the use of the more complex and realistic cell models that would otherwise be excellent candidates for investigating cellular tradeoffs [107], and even minimal mechanistic models of cells must rely on an assumed population growth property to connect the cell model to population dynamics [99]. The approach, therefore, must be simple enough to allow for sufficient mechanistic realism to capture the salient features of phenotypic tradeoffs but at the same time not lose sight of the importance of population dynamics in which adaptive evolution must take place.

1.4 A modeling framework for adaptive dynamics

To address the need for a conceptual and mathematical framework that accounts for the many factors that influence adaptive evolution, we have combined elements of ecological modeling, phenomenological mechanistic constraints, and an implementation of adaptive dynamics, to propose a method for modeling phenotypic adaptive evolution. The framework extends the foundational work of classical adaptive dynamics theory and is applied here to models of cellular resource allocation, nutrient specialization, and metabolic pathway specialization.

To implement mechanistic constraints we rely on the established theory of structured population models [60, 108], which are also by definition ecological models. In this context, we use multiple variables to track the abundance of a particular lineage, whereby the individuals that belong to a phenotypically-homogeneous lineage are further partitioned into distinct states, which communicate according to the rules and parameters of an underlying mechanistic model (Section **??**). Extensive use of the quasi steady-state approximation is utilized to allow for 'decision' points at which the flow of an individual through the state space may be diverted. Tradeoffs, such as metabolic pathway (Section **??**) selection, are implemented by parametrizing the probabilities with which a particular path in the state space is followed. The states of the population mediate different ecological interactions; therefore, the parameters, which are under adaptive evolution, that determine the state space distribution also partly determine the fitness of a lineage. We describe a way to calculate a nascent mutant lineage's invasion fitness for structured-population models based on invariant manifolds that allows for frequency-dependent selection (Section **??**).

Some assumptions in the theory of adaptive dynamics have been relaxed: first, that all phenotypic changes must be small and, second, that the phenotype space is continuous. In addition, implementation of a stochastic mutational process, which describes how the phenotype of nascent lineages deviate from that of their progenitor lineage, has also been included in the model (Chapter **??**). Removing continuity implies that it is no longer possible to work with the standard selection derivatives of adaptive dynamics theory and, as a consequence, some analytical tractability is lost; however, as will be shown, it is still possible to

recover most, if not all, of evolutionary dynamics (Sections **??** and **??**). It also becomes possible to investigate how the dynamics will change under different phenotype space discretization schemes (applied in Chapter **??** and developed in Chapter **??**). Provided the invasion fitness can be calculated, or inferred using numerical methods, we present a general way to visualize and analyze evolutionary dynamics, using principles from graph theory and discrete Markov processes (Section **??**).

We first give an analytical treatment of the adaptive dynamics (Chapter **??**), which also serves as an introduction to our approach. We then apply this analysis to the case of a well-known cellular tradeoff between biomass production and driving the cell's nutrient use pathways (Sections **??** and **??**). Using more elaborate models, but still relatively simple even when compared to minimal mechanistic replicator models [109], we then investigate the adaptive evolution of nutrient preference, in the case when multiple nutrients are present in the environment (Sections **??** and **??**), and metabolic pathway choice when a single nutrient is present in the environment (Sections **??** and **??**). For the analytically-intractable cases we have exclusively used our of numerical tools software to describe the repertoire of adaptive dynamics and ask questions about how the nature of the underlying constraints and the ecology of the cells shape evolutionary outcomes; where possible, we have compared and augmented the analytical results with the numerical methods to demonstrate their agreement and the new knowledge that is provided by a more comprehensive numerical treatment (Sections **??** and **??**).

Chapter 2

Models of adaptive dynamics in the evolution of microbial resource allocation

2.1 Introduction to chemostat models

The chemostat [38] (Figure 2.1.1) provides a laboratory-controlled idealization of nature in which the process of adaptation can be investigated (Section **??**). The models of adaptation that are presented in this chapter will all assume a chemostat-like environment, chiefly for the simplicity afforded by such an environmental context, as well as for the theoretical maturity and extensive experimental confirmation of models of continuous microbial culture in the chemostat [40].

The chemostat apparatus consists of a feed vessel, which contains all nutrients essential to growth for the microorganism of interest; notably, all of these nutrients are abundantly present except for one, limiting, nutrient. The nutrients are flowed into the growth, or culture, chamber that has been inoculated with a small sample of the organism. This chamber is kept at a constant volume, $V = l^3$ (using an arbitrary unit of length, l), by the matching of the inflow and outflow rates, which maintain a constant volumetric flow, $F = l^3/t$ (using an arbitrary unit of time, t) through the apparatus. Finally, an outflow vessel collects unused nutrients and cells, which, in an experimental context, may be sampled to provide a snapshot of the culture vessel state.

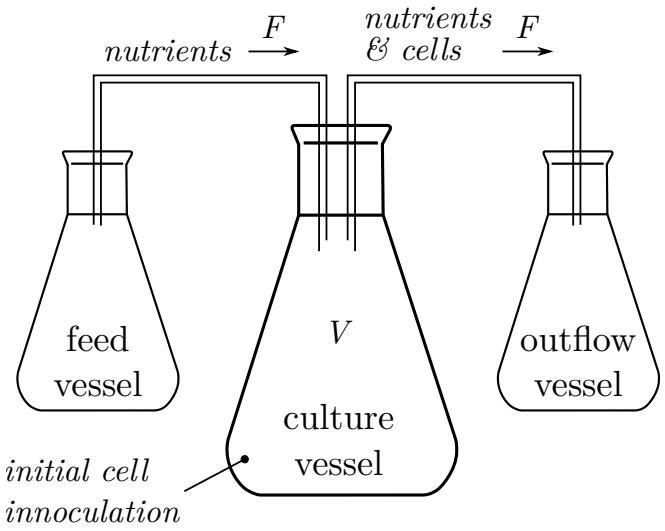


Figure 2.1.1: The chemostat vessel consists of three parts: a feed vessel, from which nutrients are supplied, a growth, or culture, chamber, and a container to collect the outflow. The culture vessel is maintained at constant volume, V, through the matching of the inflow and outflow rates, maintaining a constant volumetric flow, F, through the apparatus.

In a simple chemostat model, the limiting nutrient (hereafter referred to simply as the 'nutrient') is present at concentration $u(t)$ $mass/l^3$ and, likewise, the cells are present at concentration $x(t) mass/l^3$ at time t. The nutrient in the feed vessel has concentration $u^{(0)}$; therefore, $u^{(0)}$ $mass/t$ of nutrient is supplied per unit time.

In the absence of a growing organism, the ordinary differential equation that describes the nutrient mass is simply

$$u\dot{V}(t) = u^{(0)}F - u(t)F \tag{2.1}$$

In our models we will use the following re-parametrizations:

$$p = \frac{u^{(0)}F}{V} \quad conc./time \tag{2.2}$$

to characterize the rate of nutrient concentration influx. Moreover, we will use

$$D = \frac{F}{V} \quad 1/time \tag{2.3}$$

as the dilution rate for both cells and nutrients.

Therefore, the chemostat equation for the nutrient concentration, in the absence of any cells becomes

$$\dot{u}(t) = p - u(t)D \tag{2.4}$$

It should be noted that many alternative re-parametrizations exist, even a complete non-dimensionalization of the model [40]. We have chosen to use the above scheme as a compromise between having meaningful parameters and units, and making the expressions less cumbersome.

In this chapter we will make incremental modifications to this foundational chemostat framework to allow us to overlay and analyze models of phenotypic adaptation.

1. In Section 2.2 ('An unstructured model cellular resource allocation') the chemostat model is extended to incorporate a nutrient uptake function, which describes the interaction between cells and nutrients; moreover, we allow cells to variably allocate resources derived from metabolized nutrients toward either biomass synthesis and replication, or toward driving the metabolic pathways. The impact of different strategies of resource allocation, and the adaptation dynamics that result from the competition between strategies, will be explored.

2. In Section 2.3 ('A growth state structured model of resource allocation') the cells are classified into growth states that roughly represent (in a discrete manner) the amount of biomass accumulated by each cell. We use the growth state of an individual to track a cell's progress from birth to replication and investigate how the cellular resource allocation strategy determines adaptation dynamics in this structured model.

3. In Section 2.4 ('A structured model with intermediate cellular species') we expand the number of cell states to include intermediates that can be used to more finely control the progress of a cell through the growth state space.

4. In Section 2.5 ('A structured model with alternative metabolic pathways') we build on the model with intermediate states to allow cells to channel the nutrients taken up from the environment into one of two

alternative metabolic pathways - where each pathway is characterized by its own set of parameters (like rate and yield). We then develop the adaptation dynamics resulting from different metabolic strategies.

We have focused in this chapter, as much as possible, on a purely analytical treatment of the chemostat and adaptation models in order to develop an intuition for the salient aspects of the modeling framework with regard to adaptation dynamics. In Chapter **??** we will describe a set of numerical methods developed for working with analytically intractable models and in Chapter**??** we will apply these methods to more complicated models of adaptation.

In Chapter 4 we will use this intuition to abandon the analytic approach and explore more diverse, but analytically-intractable, models.

2.2 An unstructured model of cellular resource allocation

Following the standard chemostat setup (Section 2.1), this model consists of a single nutrient, with concentration $u(t)$, and a single cell species, with concentration $x(t)$. A well-mixed chemostat vessel, with a constant influx of nutrients (p *conc./time*) and constant outflow (D $1/time$), is inoculated with a small sample of cells at $t = 0$. Cells take up nutrients from the chemostat environment (Equation 2.5) and convert them to resources that are in turn allocated to biomass synthesis (at rate α) and toward driving the nutrient-use pathways (at rate β). A tradeoff arises from the competition between the two processes for finite resources: increasing biomass synthesis directly increases growth rate and population size, but at the expense of a decreased ability to acquire nutrients from the chemostat, which may compromise growth rate indirectly [97, 98, 99]. A schematic of the reaction channels in this system is shown in Figure 2.2.1.

Nutrient uptake is a bimolecular reaction whose rate constant is a function of the two resource allocation rates

$$k_u = k_u(\alpha, \beta) \tag{2.5}$$

We will address this rate in more detail in a few paragraphs.

We will make the assumption that the intermediate species, $\tilde{x}$, is at quasi steady state [110] with respect to the other model species. Setting, for all t:

$$\begin{aligned} \dot{\tilde{x}}(t) &= x(t)u(t)k_u - \tilde{x}(t)(\alpha + \beta) = 0 \\ \implies \tilde{x}(t) &= \frac{x(t)u(t)k_u}{\alpha + \beta} \end{aligned} \tag{2.6}$$

The quasi steady state is frequently encountered in biochemical kinetics. Its application to resource-consumer models is predicated on several assumptions. First, we are treating both microbes and nutrients as freely-diffusing molecules in a well-mixed environment. For large bio-reactors and chemostats, this is an acceptable practice [40], in part because of the following assumption. Second, we are assuming large numbers of both nutrients and cells – sufficiently large so that cells are effectively saturated with nutrients. Finally, we are accepting

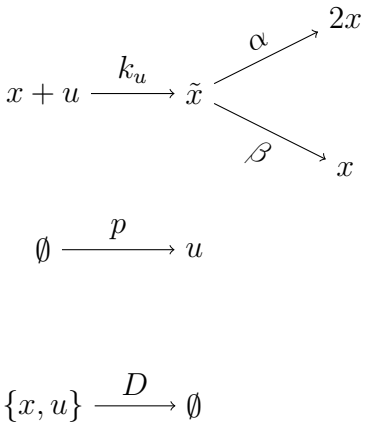


Figure 2.2.1: Schematic for the simple, unstructured, population model. There are three reaction channels: nutrient uptake, nutrient influx, and dilution (of cells and nutrients); in addition, there are two 'nested' reaction channels that result from resource allocation: biomass synthesis/replication (α) and nutrient-use pathway maintenance (β). These operate at quasi steady state according to Equation 2.6.

the existence of a fast timescale on which an initial transient exists where the intermediate species concentrations might temporarily change – transiently violating $\dot{\tilde{x}}(t) = 0$, and potentially invalidating the approximation. For our purposes, however, the initial transient is not critical, because we are working with models where we are always pairing a large (fast) reaction rate with a small (slow) reaction rate. The ratio of these parameters determines the fast timescale and shrinks the duration of the initial transient, relative to the characteristic time of cell growth, and ensures that the error from the approximation is small [110].

In other words, we have stated that the resource-allocation "decision" is effectively instantaneous. The probability of allocating one unit of nutrient toward biomass synthesis, then, is

$$P_{\text{biomass}} = \frac{\alpha}{\alpha + \beta} = r_x \tag{2.7}$$

and toward the nutrient-use pathways

$$P_{\text{nutrient}} = \frac{\beta}{\alpha + \beta} = 1 - r_x \tag{2.8}$$

$r_x \in [0, 1]$ is the resource allocation parameter and characterizes the phenotype of the cell population. Unlike other model parameters, which are fixed, we will

later investigate the adaptive dynamics with respect to the resource allocation parameter.

We will also make the simplifying assumption that biomass synthesis and driving of the nutrient-use pathways are not rate-limiting reactions. This assumption will be relaxed in later models by the introduction of more intermediate species.

In light of the quasi steady state assumption, then, the nutrient-uptake rate, k_u, becomes a function of the resource allocation parameter

$$k_u = k_u(r_x) \tag{2.9}$$

We will not assume a specific function for the nutrient uptake; nevertheless, a monotone decreasing function makes the most sense for k_u. An inverted Hill equation will be used later to illustrate examples.

The rate equations for the two dynamical variables constitute a simple two-dimensional system of ordinary differential equations (2.10). This is the *resident* model - in keeping with the naming conventions from the theory of adaptive dynamics [68, 73] - and follows closely the traditional form of the equations describing the continuous culture of microbial populations [40, 111].

$$\begin{aligned} \dot{x} &= xuk_u(r_x)r_x - xD \\ \dot{u} &= p - uD - uxk_u(r_x) \end{aligned} \tag{2.10}$$

p	nutrient influx rate
D	dilution rate
r_x	resource allocation parameter
$k_u(r_x)$	nutrient uptake function

Table 2.1: Parameters of the unstructured model. *Top*: system parameters. *Bottom*: adaptation parameters

We will primarily be concerned with the fixed points and stationary behavior of the system. A typical trajectory is shown in Figure 2.2.2.

There are two fixed points in this system where the derivatives vanish:

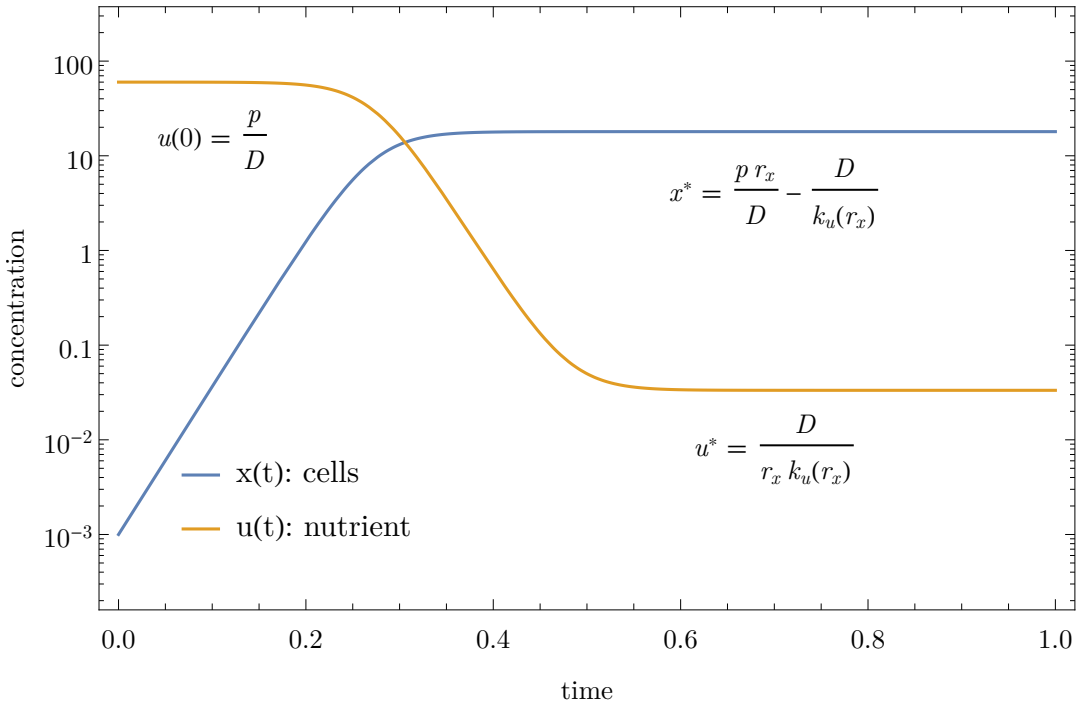


Figure 2.2.2: Simulation of the unstructured model showing cell concentration, $x(t)$, and nutrient concentration, $u(t)$, trajectories. Annotations show the steady states and nutrient initial value (which is the nutrient steady state in the absence of cells). The initial value of cell population is any small value (here 10^{-3}). The axes are in arbitrary units of concentration (c) and time (t). The parameters for this example are: $p = 1.2ct^{-1}, D = 0.02t^{-1}, k_u(r_x) = 2(ct)^{-1}, r_x = 0.3$.

$$\begin{aligned} x^* &= 0 \\ u^* &= \frac{p}{D} \end{aligned} \tag{2.11}$$

and

$$\begin{aligned} x^* &= -\frac{D}{k_u(r_x)} + \frac{pr_x}{D} \\ u^* &= \frac{D}{k_u(r_x)r_x} \end{aligned} \tag{2.12}$$

The two steady states correspond to the extinction (2.11) and survival (2.12) events. The stability of the steady states can be assessed using linear stability analysis. The Jacobian matrix of the system is

$$\boldsymbol{J}_R = \begin{bmatrix} -D + k_u(r_x)r_x u & k_u(r_x)r_x u \\ -k_u(r_x)u & -D - k_u(r_x)x \end{bmatrix} \tag{2.13}$$

which has eigenvalues:

$$\begin{aligned} \lambda_1 &= -D \\ \lambda_2 &= -D + k_u(r_x)r_x u - k_u(r_x)x \end{aligned} \tag{2.14}$$

The second eigenvalue determines stability, since the dilution rate is always positive. Evaluating the expression for λ_2 at the extinction steady state gives the extinction condition:

$$k_u(r_x)r_x < \frac{D^2}{p} \tag{2.15}$$

where the extinction point is stable. A cell population characterized by resource allocation parameter r_x and nutrient uptake rate $k_u(r_x)$ will not fix in the chemostat (Equation 2.15) and will be completely washed out. The survival condition is simply the reverse:

$$k_u(r_x)r_x > \frac{D^2}{p} \tag{2.16}$$

So the system always has a single attractor. Intuitively, the above inequality implies that 'harsher' environments, with higher dilution rates and/or lower nutrient availability, limit the viable phenotypes (r_x) by requiring a higher product of the nutrient uptake rate and biomass synthesis/replication.

A more intuitive way to interpret Equation 2.16 is to notice that survival simply requires that the concentration of cells initially grows when added to a chemostat that has reached steady state in the absence of cells. Using the nutrient steady state from Equation 2.11 in the rate equations:

$$\begin{aligned} \dot{x} &= xuk_u(r_x)r_x - xD \\ &= x\left[uk_u(r_x)r_x - D\right]_{u=u^*=p/D} \\ &= x\left[k_u(r_x)r_x - \frac{D^2}{p}\right] \end{aligned} \tag{2.17}$$

where the square brackets contain the extinction (Equation 2.15, when negative) and survival (Equation 2.16, when positive) conditions. The conditioning on the state of the environmental variables (nutrient) is a recurring approach that we will encounter often in this chapter.

For the remainder of this section we will assume that we are operating in the survival parameter regime - i.e. dilution rates are small relative to the nutrient influx rate and the nutrient uptake rate is adequate to support growth.

To analyze the adaptive dynamics of the resource allocation parameter we consider the augmented *resident-mutant* model:

$$\begin{aligned} \dot{x} &= xuk_u(r_x)r_x - xD \\ \dot{u} &= p - uD - u(xk_u(r_x) + yk_u(r_y)) \\ \dot{y} &= yuk_u(r_y)r_y - yD \end{aligned} \tag{2.18}$$

Where the mutant lineage has concentration $y(t)$ and is characterized by the resource allocation parameter $r_y \neq r_x$. The mutant lineage is initially rare relative to the resident lineage, which has initial concentration equal to the steady state of the *resident* model (Equation 2.12). Therefore, for the *resident-mutant* model, we have the initial condition:

$$\begin{aligned} x(t=0) &= x^{*resident*} \\ u(t=0) &= u^{*resident*} \\ y(t=0) &= \epsilon \end{aligned} \tag{2.19}$$

for some vanishingly small value of $\epsilon \to 0$.

The fate of the nascent mutant lineage can be determined using stability analysis. The Jacobian of the augmented model

$$\boldsymbol{J}_{RM} = \begin{bmatrix} \boldsymbol{J}_R & \cdots \\ \mathbf{0} & -D + k_u(r_y)r_y u \end{bmatrix} \tag{2.20}$$

Where $\boldsymbol{J}_R$ is the *resident* model's (Equation 2.10) Jacobian matrix (Equation 2.13). Because of the matrix's block-triangular structure, its eigenvalues consist of the union of the eigenvalues of $\boldsymbol{J}_R$ and the eigenvalues of the 1x1 matrix in the

lower-right corner [112]. Moreover, the *resident* model's survival steady state (Equation 2.12) is also a fixed point for the *resident-mutant* model.

The mutant lineage will invade if the point $(x, u, y) = (x^{*resident*}, u^{*resident*}, 0)$ is unstable. The eigenvalues from J_R at this point are already negative - by definition, since we have assumed that the resident lineage reaches (dynamically-stable) survival before the mutant lineage emerges; therefore, we only need to evaluate the third eigenvalue:

$$\begin{aligned} \lambda_3 = \lambda(r_x, r_y) &= -D + k_u(r_y) r_y u^{*resident*} \\ &= D\left(-1 + \frac{k_u(r_y) r_y}{k_u(r_x) r_x}\right) \end{aligned} \tag{2.21}$$

The $\lambda(r_x, r_y)$ eigenvalue is known as the *invasion eigenvalue* and, when positive, gives the initial exponential growth rate of the nascent mutant population in the chemostat environment modified by the resident lineage. Indeed, the invasion eigenvalue is equal to the per-capita growth rate, $\dot{y}(t)/y(t)$, of the mutant lineage in the early phase of the invasion [73]. Such a quantity has traditionally been referred to as the invasion fitness [69, 73]. It is important to note that the success or failure of the mutant in the chemostat environment is contingent on the presence and phenotype (r_x) of the resident lineage - that is, the mutant lineage's growth rate is frequency-dependent.

Equation 2.21 gives the invasion condition

$$\boxed{k_u(r_y) r_y > k_u(r_x) r_x} \tag{2.22}$$

The invasion-implies-substitution theorem [73] applies here and guarantees that a mutant lineage that invades successfully (i.e. does not become extinct) will drive the resident lineage to extinction. It follows that the mutant lineage will reach steady state

$$y = y^{*mutant*} = -\frac{D}{k_u(r_y)} + \frac{p r_y}{D} \tag{2.23}$$

which is the same steady state from the previous *resident* model with $r_x \to r_y$, since there is only one extant lineage after the resident extinction. The chemostat environment will now be determined entirely by the mutant lineage, so the system

will have steady state

$$(x, u, y) = (0, u^{*mutant*}, y^{*mutant*}) = (0, \frac{D}{k_u(r_y)r_y}, -\frac{D}{k_u(r_y)} + \frac{pr_y}{D}) \tag{2.24}$$

The invasion condition (Equation 2.22) can be understood in terms of a lineage's ability to extract nutrients from the environment. In particular, since

$$\begin{aligned} u^{*resident*} &= \frac{D}{k_u(r_x)(r_x)} \\ u^{*mutant*} &= \frac{D}{k_u(r_y)(r_y)} \end{aligned} \tag{2.25}$$

are the nutrient steady states of the resident-only and mutant-only models, we can rewrite the invasion condition as

$$u^{*resident*} > u^{*mutant*} \tag{2.26}$$

Therefore a nascent mutant lineage, with phenotype r_y, will invade and drive a resident lineage, with phenotype r_x, to extinction if it can further decrease the concentration of nutrients in the chemostat. This is in accordance with the 'R^*' concept of competitive exclusion (exclusion of all lineages by the lineage with the fastest growth rate [42]). A lineage will grow and extract nutrients from the environment to a concentration where its growth rate rate exactly matches the rate of removal (D). The population cannot grow below this nutrient concentration. When a mutant lineage emerges, whose 'break-even' nutrient concentration is lower than that of the resident, the nutrient concentration in the chemostat will decrease below the point that supports a positive population growth for the resident lineage. The resident lineage will then asymptotically approach extinction as the mutant lineage reaches its own steady state.

The implications of these results demonstrate the utility of resource-based models over the models of Lotka, Volterra, and Malthus [36, 37]. In the latter, competitive interactions are phenomenological, with no regard to the mechanisms by which a competitor affects the growth rate of a resident species; as a result, the competition parameters can only be measured when two competitor species are grown together. This is in contrast to adaptive dynamics and resource-consumer models, where the relevant parameters (in this case $k_u(r_x)(r_x)$, or simply u_x^*) can

be determined in advance of competition, and verified experimentally [46, 47].

The invasion condition (Equation 2.22), therefore, is sufficient to predict the outcome of any invasion event between any resident and mutant lineage, and can be used to construct an evolutionary process; for this section, however, we will continue beyond the invasion condition to derive a smooth dynamical process for the trajectory of the resource allocation parameter in evolutionary time.

Expanding the invasion eigenvalue $\lambda(r_x, r_y)$ in Taylor series around $r_y \approx r_x$

$$\lambda(r_x, r_y) = \lambda(r_x, r_x) + \frac{\partial}{\partial r_y}\lambda(r_x, r_y)\Big|_{r_y=r_x} + O(r_y)^2 \tag{2.27}$$

The invasion eigenvalue $\lambda(r_x, r_x) = 0$ because invasion by a lineage with the same resource allocation parameter is neutrally stable - i.e. the nascent lineage will neither grow or decline deterministically (and would almost certainly be lost due to accidental extinction). The partial derivative

$$\frac{\partial}{\partial r_y}\lambda(r_x, r_y)\Big|_{r_y=r_x} = D\left(\frac{1}{r_x} + \frac{1}{k_u(r_x)}\frac{\partial}{\partial r_y}k_u(r_y)\Big|_{r_y=r_x}\right) \tag{2.28}$$

is known as the selection derivative [73]. We can write the "canonical" equation of adaptive dynamics by incorporating details about the mutation process, which determines the rate at which beneficial mutations arise, and about the stochastic birth-death process, which decreases the rate of evolution due to accidental extinctions of beneficial mutations [73, 113]. The canonical equation is:

$$\begin{aligned}\frac{dr_x}{dt} &= \frac{1}{2}\mu(r_x)\sigma^2(r_x)x^{*resident*}\frac{\partial}{\partial r_y}\lambda(r_x, r_y)\Big|_{r_y=r_x} \\ &= \frac{1}{2}\mu(r_x)\sigma^2(r_x)x^{*resident*}D\left(\frac{1}{r_x} + \frac{1}{k_u(r_x)}\frac{\partial}{\partial r_y}k_u(r_y)\Big|_{r_y=r_x}\right)\end{aligned} \tag{2.29}$$

Here μ is the mutation rate (the rate at which daughter cells appear that have an adaptation/phenotype parameter that is different from that of its immediate ancestor) and σ is the standard deviation of the mutational step (a measure of how different the mutant parameter is relative to the ancestral one), both of which may be functions of the resource allocation parameter. The exact form of equation 2.29 depends, of course, on the nutrient-use function and its derivative.

Equation 2.29 describes the evolution of the resource allocation parameter of

the resident lineage (r_x) through evolutionary time under the sequential invasion of nascent mutant lineages - taken to the limit of a smooth process under the previously mentioned assumptions.

Following is a brief overview of the derivation of the canonical equation, following Dercole and Rinaldi [73]. We begin by noting that evolutionary change is a stochastic process with two major sources of stochasticity: mutation and demography. Therefore, a deterministic description of the rate of change of a phenotypic parameter over evolutionary time must be interpreted as the average evolutionary change over all possible realizations of stochastic change. In other words

$$\frac{dr_x}{dt} = \lim_{dt\to 0} \frac{E[r_x(t+dt) - r_x(t)]}{dt} \tag{2.30}$$

where the expectation is taken with respect to the stochastic processes involved. If we denote $P(r_x, r_{x'})dr_{x'}$ as the probability that a population with phenotype r_x at time t will be characterized by phenotypes between $r_{x'}$ and $r_{x'} + dr_{x'}$ at time $t + dt$, then Equation 2.30 becomes

$$\frac{dr_x}{dt} = \lim_{dt\to 0} \frac{1}{dt} \int_{-\infty}^{\infty} (r_{x'} - r_x) P(r_x, r_{x'}) dr_{x'} \tag{2.31}$$

The probability term can be calculated from the product of three probabilities.

Mutation probability The probability of a mutation in the evolutionary time interval $[t, t + dt]$ can be derived from the birth process at the demographic time scale. Suppose that dt on the evolutionary time scale corresponds to $\frac{dt}{\epsilon}$ of demographic time, for some small ϵ that separates the time scales. To first order, each of the $1/\epsilon$ demographic time intervals has a probability of generating a mutant cell

$$\epsilon \cdot \mu(r_x) \cdot \mathcal{B}(r_x, x^*) \cdot x^* dt + \mathcal{O}(dt^2)$$

Here $\mu(r_x)$ is the mutation rate (per birth), which may be a function of the phenotypic parameter. $\mathcal{B}(r_x, x^*)$ is the per-capita birth rate of the lineage with phenotype r_x, evaluated at the steady state of a population containing this lineage, x^*. The probability that a single mutation occurs in any one of the demographic intervals, and not in any other, is then

$$P_m(r_x) = \mu(r_x) \cdot \mathcal{B}(r_x, x^*) \cdot c^* dt + \mathcal{O}(dt^2) \tag{2.32}$$

Mutational step We will further assume that the mutational step, $(r_{x'} - r_x)$ is drawn from a symmetric probability distribution centered at zero. We will require that the standard deviation of the distribution vanishes as $\epsilon \to 0$. We can then write

$$P'(r_x; r_{x'} - r_x) = \frac{1}{\epsilon} D\left(r_x; \frac{r_{x'} - r_x}{\epsilon}\right)$$

using a suitable rescaling of some parameter distribution D with standard deviation σ, as the probability distribution of the mutational step $r_{x'} - r_x$ when the population has a single lineage with phenotype r_x. The variance of the mutational step is then

$$E[(r_{x'} - r_x)^2] = \epsilon^2 \sigma^2(r_x)$$

Probability of accidental extinction A mutant phenotype may be lost by stochastic extinction when it is initially present at very low numbers. Following the Moran process [6] we can write the probability that the mutant survives (and fixes) in the environment starting with a single individual as

$$P_s = \frac{1 - r^{-1}}{1 - r^{-N}}$$

The selective advantage, r, of the mutant can be written as the ratio of the per-capita birth to death rate of the mutant

$$r = \frac{\mathcal{B}(r_x, r_{x'}, x^*)}{\mathcal{D}(r_x, r_{x'}, x^*)}$$

evaluated at the steady state of resident concentration, x^*. Taken to the limit of infinitely large population, $N \to \infty$, the survival probability is

$$P_s(r_x, r_{x'}) = \frac{\lambda(r_x, r_{x'})}{\mathcal{B}(r_x, r_{x'}, x^*)} \tag{2.33}$$

since the invasion eigenvalue is the difference between the per-capita birth and death rate of the mutant evaluated at the resident population's steady state:

$\lambda(r_x, r_{x'}) = \mathcal{B}(r_x, r_{x'}, x^*) - \mathcal{D}(r_x, r_{x'}, x^*)$. Note that the above holds only for $r > 1$, which implies that $\lambda(r_x, r_{x'}) > 0$.

Multiplying the three probabilities gives

$$\begin{aligned} P(r_x, r_{x'})dr_{x'} &= P_m(r_x) \cdot P'(r_x, r_{x'}) \cdot P_s(r_x, r_{x'})dr_{x'} \\ &= \mu(r_x) \cdot \mathcal{B}(r_x, x^*) \cdot P'(r_x; r_{x'} - r_x) \cdot P_s(r_x, r_{x'})dr_{x'}dt \end{aligned} \tag{2.34}$$

Substituting the product in Equation 2.31 we get

$$\frac{dr_{x'}}{dt} = \mu(r_x) \cdot \mathcal{B}(r_x, x^*) \cdot x^* \int_{-\infty}^{\infty} dr_{x'}(r_{x'} - r_x) \cdot P'(r_x; r_{x'} - r_x) \cdot \frac{\lambda(r_x, r_{x'})}{\mathcal{B}(r_x, r_{x'}, x^*)}$$

which, after the first order expansion of $\lambda(r_x, r_{x'})$ (see Equation 2.27) becomes

$$\frac{dr_{x'}}{dt} = \mu(r_x) \cdot x^* \cdot \frac{\partial}{\partial r_{x'}} \lambda(r_x, r_{x'})\Big|_{r_{x'}=r_x} \int_{r_x}^{\infty} dr_{x'}(r_{x'} - r_x)^2 \cdot P'(r_x; r_{x'} - r_x)$$

Due to the symmetry of the mutation distribution, and since we are only evaluating the half of the distribution for which the selection derivative is positive, the integral is simply half of the variance of the mutation distribution, which we'd earlier set to σ^2. Therefore, we finally get

$$\frac{dr_{x'}}{dt} = \frac{1}{2}\epsilon^2 \mu(r_x)\sigma^2(r_x) \cdot x^* \cdot \frac{\partial}{\partial r_{x'}} \lambda(r_x, r_{x'})\Big|_{r_x=r_{x'}} \tag{2.35}$$

which can be rescaled by defining a new time variable to remove the ϵ factor, thus yielding Equation 2.29.

2.2.1 Applications and adaptive dynamics

A monotone decreasing function is a sensible choice for the nutrient-use rate, k_u, in terms of resources invested toward driving the nutrient uptake pathways, $k_u(r_x)$. Holling [114] originally proposed functional forms for predator-prey interactions by considering aspects of predation like 'searching' and 'handling'. Michaelis-Menten functional responses for microbial populations have previously

been successfully used to study competition (for example, in [115]); keeping with this tradition, we will use an inverted Hill-type equation to incorporate resource allocation into the nutrient use rate function:

$$k_u(r_x) = k_{min} + (k_{max} - k_{min})\frac{(1 - r_x)^\eta}{(1 + K)^\eta + (1 - r_x)^\eta} \tag{2.36}$$

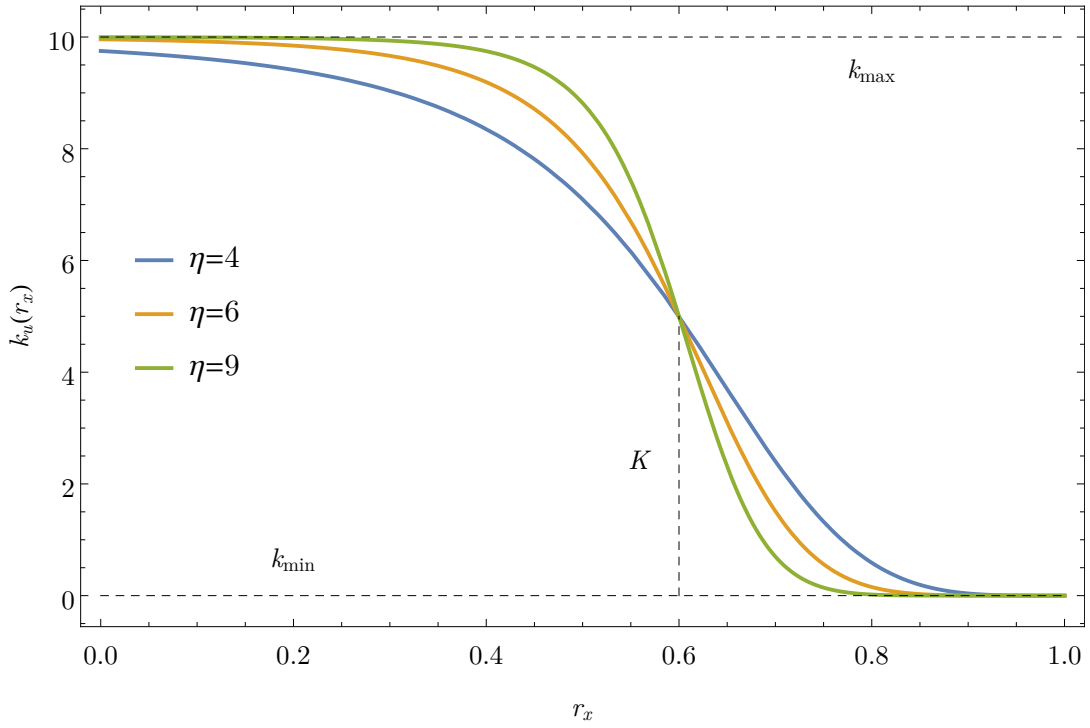


Figure 2.2.3: The inverted Hill nutrient uptake rate (equation 2.36), shown with three different Hill coefficients (η), is a monotone decreasing function of the fraction of resource allocated toward biomass synthesis $(1 - r_x)$. The vertical axis is in arbitrary units of $\frac{1}{conc.\cdot time}$. Parameters used for this example: $k_{min} = 0\ (ct)^{-1}, k_{max} = 10\ (ct)^{-1}, K = 0.6$.

Using the inverted Hill equation as the nutrient uptake rate function, and its derivative with respect to r_x, in equation 2.29, we can evaluate the (local) evolutionary dynamics of the resource allocation parameter, r_x. Figure 2.2.4 shows the magnitude and direction of the resource allocation parameter's derivative in evolutionary time in the neighborhood of a resident lineage with phenotype r_x for example parameters.

The evolving system reaches steady state when the derivative vanishes. In addition, the system is stationary when the derivative is positive at the upper boundary ($r_x = 1$), or when the derivative is negative at the lower boundary, $r_x = 0$,

because of the bounded $r_x \in [0, 1]$. These fixed points of the system are known as the (locally) evolutionarily singular points [63, 65, 66, 68] because no mutant lineage - in this case, with phenotype close to that of the ancestral resident's phenotype - can invade. Figure 2.2.4 shows the general form of the derivative under the assumption that $k_{max} > k_{min}$.

The system has one or two stationary points:

1. If the stable steady state point (where the derivative vanishes) exists, it is an evolutionarily singular strategy. In addition, the unstable steady state point always exists if the stable steady state point exists, and is the boundary between the basin of attraction for the evolutionarily singular strategy at $r_x = 1$ and the evolutionarily singular strategy at the stable steady state point (Figure 2.2.4, green and orange lines).

2. If the stable steady state point does not exist, the upper boundary, $r_x = 1$ is an evolutionarily singular strategy because the derivative is positive. All values of r_x will converge to this point (Figure 2.2.4, blue line).

Substituting the resident model's (survival) steady state (equation 2.12 in Equation 2.29, and omitting the strictly positive mutation rate (μ) and mutational size (σ) terms, the equation for the evolution of r_x becomes

$$\frac{dr_x}{dt} \propto \frac{\left\{ r_x \frac{\partial}{\partial r_x}(r_x) + k_u(r_x) \right\} (p r_x k_u(r_x) - D^2)}{r_x k_u(r_x)^2} \tag{2.37}$$

Only the term in curly brackets may become negative. $(p r_x k_u(r_x) - D^2)$ is positive by assumption since it is the condition for the extinction steady state point of the resident model to be unstable; likewise, the denominator is always positive since $r_x \in [0, 1]$ and the nutrient-use rate function is non-negative. There is no closed form for the solution of $r_x \frac{\partial}{\partial r_x}(r_x) + k_u(r_x) = 0$ since this involves a polynomial of order η (the Hill coefficient) that is not necessarily an integer. Therefore it is not always possible to determine analytically when the adaptive dynamics system has one or two stationary points.

Beyond the local dynamics of Equation 2.29 we can use the invasion condition (Equation 2.22) to examine the global behavior of the system when we remove the restriction for small mutations.

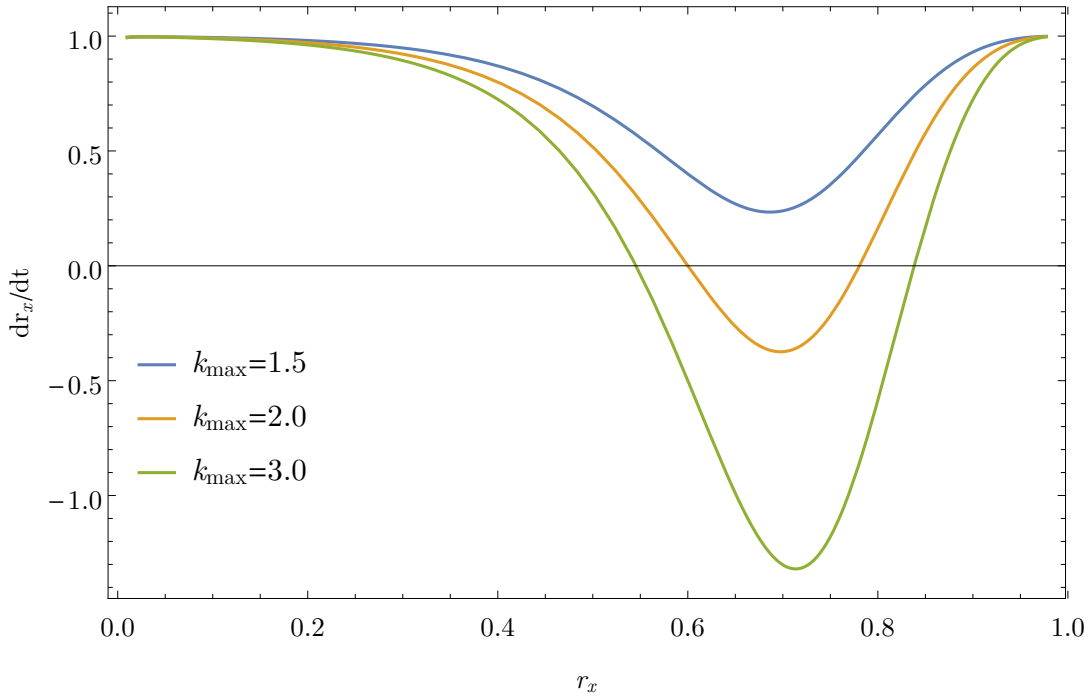


Figure 2.2.4: Local direction of change of the resource allocation parameter r_x. In addition to the steady state points where the derivative vanishes (horizontal axis crossing), the boundaries $r_x = 0$ and $r_x = 1$ are also stationary points when $dr_x/dt < 0$ and $dr_x/dt > 0$, respectively, because of the bounded parameter. Parameters used: $\eta = 4, K = 0.6, D = 10^{-2}\ 1/t, p = 1\ ct^{-1}$, and $k_{min} = 1\ (ct)^{-1}$. The mutation rate and size of mutational step were constant for all values of r_x and would not affect the overall shape of dynamics; therefore $\frac{1}{2}\mu(r_x)\sigma^2(r_x) = 1$ was used for this figure.

Figure 2.2.5 shows the *pairwise invasibility plots* [73] using the same example parameters used to visualize the local evolutionary dynamics in Figure 2.2.4. The invasion condition provides a map for all possible evolutionary trajectories of the resource allocation parameter r_x since it predicts the invasion outcome of any mutant lineage in the background of any resident lineage (examined more fully in Chapter 4). This approach is more qualitative, and less rigorous, than the smooth dynamical process provided by Equation 2.29; however, it may be the only way to fully describe the entirety of the model's evolutionary behavior, and, as we will see in Chapter 4, it provides the basis for a more complete analysis of global adaptive dynamics. We will use the invasion condition exclusively for the following sections in this chapter.

The simplicity of the invasion condition means that the invasion fitness of a particular lineage can be written as

$$f_I(r_x) = [k_u(r_x)r_x]^{-1} \tag{2.38}$$

In the sense that lineage r_y will invade and drive a resident lineage r_x to extinction if $f_I(r_y) > f_I(r_x)$. This is in agreement with the competitive exclusion principle [40, 45] (Section **??**).

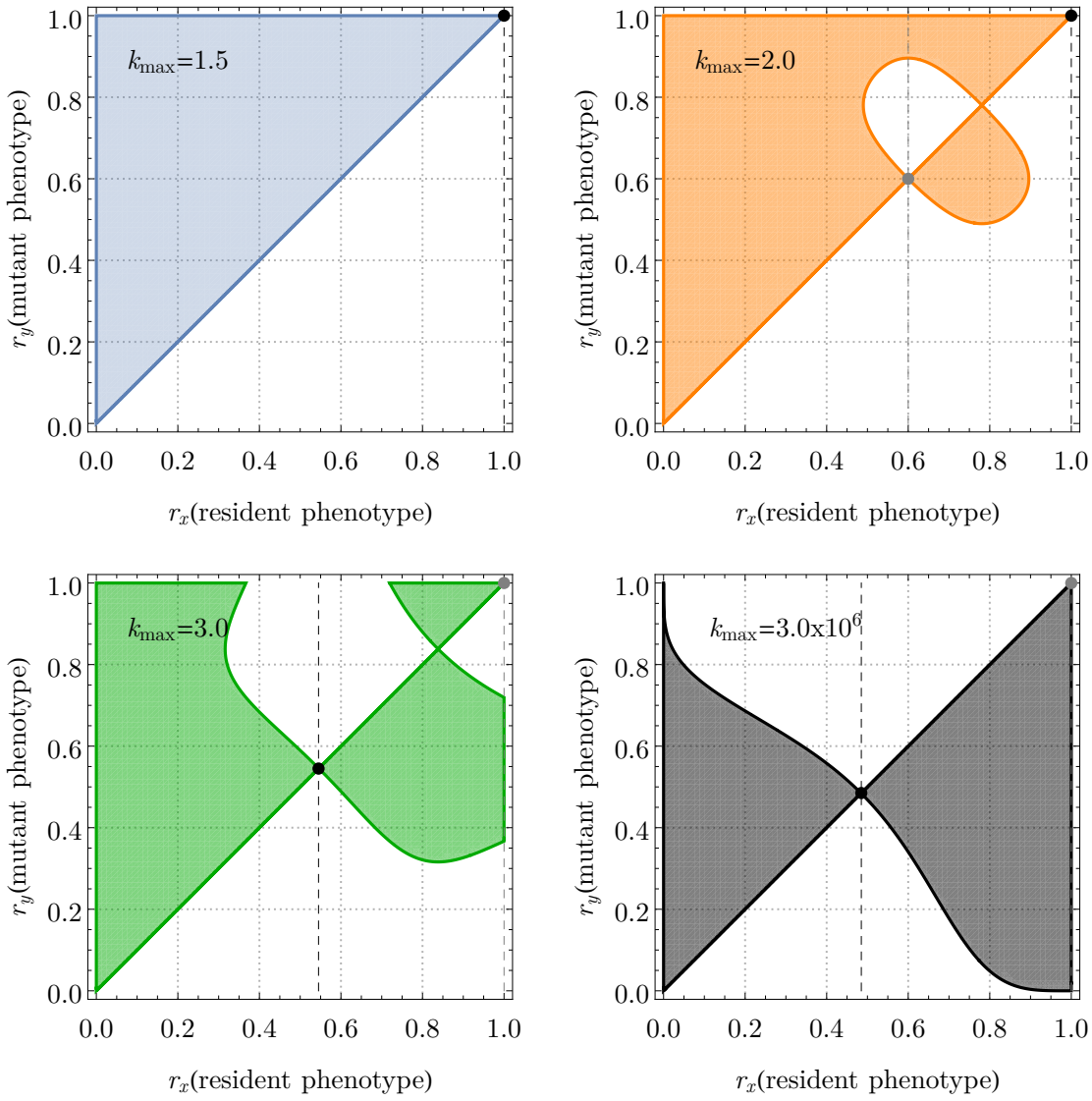


Figure 2.2.5: Pairwise invasibility plots showing the outcome of an invasion event between every resident lineage with resource allocation phenotype r_x and a nascent mutant lineage with resource allocation phenotype r_y. Shaded regions show the pairs of resident-mutant values where the mutant invades and drives the resident to extinction. Black circles show globally evolutionarily singular strategies - that is, values of r_x where no $r_y \in (0, 1)$ can invade. Gray circles show locally evolutionarily singular strategies - in other words, values of r_x where no $r_y \approx r_x$ can invade. The figures have been color-coded to correspond to the smooth (local) adaptive dynamics shown in Figure 2.2.4. The frequency-dependent nature of analyzing adaptive dynamics in this manner is more pronounced when visualizing the global evolutionary behavior of the system. Parameters used: $\eta = 4, K = 0.6, D = 10^{-2}\ 1/t, p = 1\ ct^{-1}$, and $k_{min} = 1\ (ct)^{-1}$.

In this section we have laid the basis for formulating microbial competition models where cells allocate a finite resource between competing processes - implementing a phenotype-determined strategy. We have also seen how the models are analyzed with principles from adaptive dynamics theory in both a local and a global scope, and we have visualized the adaptation process for an example parameter set. A noteworthy observation - that in large part motivated the numerical methods described in Chapters **??** and **??** - is that the magnitude and statistics of the phenotype mutations, together with the initial (ancestral) value of the phenotype, are crucial for predicting the long-term outcome of adaptation (Figure 2.2.5).

2.3 A growth state structured model of resource allocation

We continue the modification of the basic chemostat model by formulating a structured model where the consumers (cells) are grouped into several variables. The model is similar to the previous, unstructured, model (Equation 2.10) in all respects except for the presence of n cell (growth) states instead of just one. The cell states are indexed from 1 to n. Nutrient uptake, followed by biomass synthesis (at rate α), forces the cell to transition to the next-highest state. If the cell is at state n, which we take as a growth threshold, then two daughter cells are produced in the first (basal) state. As with the unstructured model, resources are also allocated toward driving the nutrient-use pathways (at rate β). Nutrients are added into the chemostat vessel at a constant rate, $p\ ct^{-1}$, and both nutrients and cells are washed out at a constant rate, $D\ t^{-1}$. Figure 2.3.1 shows the model's reaction channels.

Conceptually, this model will allow us to more finely partition a population of cells using a structure that, in this case, distinguishes between cells according to their biomass. In later models, we will incorporate more elaborate modes of transition between growth states that are in part determined by the population's resource allocation strategy. From a technical standpoint, we will see how structured models can be mapped to the simpler unstructured models.

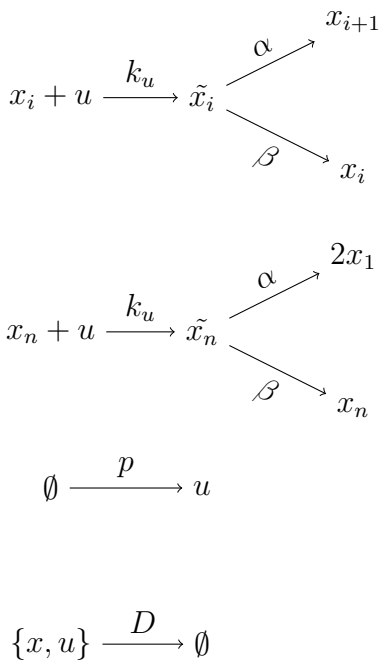


Figure 2.3.1: Schematic for the structure population model reaction channels. There are $n+1$ dynamical variables in the model: n x_i cell states, and one nutrient variable, u.

Making the assumption that the intermediate species, $\tilde{x}_i$, are at quasi steady state with respect to the rest of the system, and again setting $r_x = \frac{\alpha}{\alpha+\beta}$, the resident model has equations:

$$\begin{aligned} \dot{x}_1 &= uk_u(r_x)r_x\left(-x_1 + 2x_n\right) - x_1 D \\ \dot{x}_i &= uk_u(r_x)r_x\left(-x_i + x_{i-1}\right) - x_i D \quad \text{for } 1 < i \leq n \\ \dot{u} &= p - uD - uk_u(r_x)\sum_{i=1}^{n} x_i \end{aligned} \tag{2.39}$$

Or, in matrix form,

$$\begin{aligned} \dot{\mathbf{x}} &= uk_u(r_x)r_x\mathbf{A}\mathbf{x} - \mathbf{x}D \\ \dot{u} &= p - uD - uk_u(r_x)\mathbf{1}\cdot\mathbf{x} \end{aligned} \tag{2.40}$$

with transition matrix

$$\boldsymbol{A} = \begin{bmatrix} -1 & 0 & 0 & \dots & 2 \\ 1 & -1 & 0 & \dots & 0 \\ 0 & 1 & -1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & -1 \end{bmatrix} \tag{2.41}$$

First, note that the per-capita growth rates are not trivial even at steady state, unlike the unstructured model (Equation 2.21); for example:

$$\frac{\dot{x}_i(t)}{x_i(t)} = uk_u(r_x)r_x \left(-1 + \frac{x_{i-1}(t)}{x_i(t)} \right) - D \tag{2.42}$$

requires the distribution of cell state concentrations at time t. Recall that the per-capita growth rate is equal to the invasion eigenvalue in the (unstructured) resident-mutant model (Equation 2.21); in this model, however, there are multiple states and therefore multiple per-capita growth rates. We will now derive expressions for the steady state distribution, which lies on an invariant manifold, and we will use these expressions to calculate the steady state of the resident model, the invasion eigenvalues of the resident-mutant model, and the invasion condition. We use the operating definition of an invariant manifold in the discussion that follows: a continuously and smoothly parametrizable manifold that is also an invariant set (of points in the dynamical system which are mapped onto other points in the set by the same operator).

Because of the step-like state progression we have, for $i > 1$ at steady state (indicated by the asterisk superscript):

$$\frac{x_{i-1}^*}{x_i^*} = \frac{u^* k_u(r_x)r_x + D}{u^* k_u(r_x)r_x} = \frac{h + D}{h} \tag{2.43}$$

where we have defined

$$h = u^* k_u(r_x)r_x \tag{2.44}$$

We can chain the ratios together to write

$$\frac{x_i^*}{x_n^*} = \left(\frac{x_{n-1}^*}{x_n^*} \right) \left(\frac{x_{n-2}^*}{x_{n-1}^*} \right) \dots \left(\frac{x_i^*}{x_{i+1}^*} \right) \tag{2.45}$$

which will be useful later.

All state derivatives vanish at steady state, therefore their sum should also vanish. Letting $X = \sum_{i=1}^{n} x_i$, we have

$$\begin{aligned} \dot{X} = 0 &= u^* k_u(r_x) r_x \left[\cancel{-x_1} + 2x_n - \cancel{x_2} + \cancel{x_1} - \cdots - \cancel{x_n} + \cancel{x_{n-1}} \right] - DX \\ \implies \frac{h}{D} &= \sum_{i=1}^{n} \frac{x_i^*}{x_n^*} \\ \frac{h}{D} &= \sum_{i=1}^{n} \left(\frac{h+D}{h} \right)^{n-i} \end{aligned} \tag{2.46}$$

The RHS term is a geometric series and the expression simplifies to

$$h \left(\frac{D+h}{h} \right)^{n} = 2h \tag{2.47}$$

which gives the simple expression:

$$\boxed{h = \frac{D}{2^{1/n} - 1}} \tag{2.48}$$

Substituting this expression for h in the steady state ratios (equation 2.43) we obtain simple expressions for the steady state distribution:

$$\begin{aligned} \frac{x_{i-1}^*}{x_i^*} &= 2^{1/n} \\ \frac{x_i^*}{x_n^*} &= 2^{1-i/n} \end{aligned} \tag{2.49}$$

The nutrient steady state can also be simplified:

$$u^* = \frac{D}{(2^{1/n} - 1) k_u(r_x) r_x} \tag{2.50}$$

Compare this to the nutrient steady state for the unstructured model, $u^* = D/(k_u(r_x)r_x)$. When cells require more nutrients before they can replicate, the nutrient steady state in the chemostat increases since the size of the population that can be supported becomes smaller (see Equation 2.53) and more unused nutrients remain in the chemostat.

We can also write an expression for the steady state of the cell states. First:

$$\begin{aligned} \dot{u} = 0 &= p - u^* D - u^* k_u(r_x) \frac{h}{D} x_n^* \\ \implies x_n^* &= \frac{D(p - u^* D)}{u^* k_u(r_x) h} \end{aligned} \tag{2.51}$$

and use this in

$$\begin{aligned} \dot{x}_1 = 0 &= h x_1^* + 2h x_n^* - x_1^* D \\ x_1^* &= 2 \left(\frac{h}{D - h} \right) x_n^* \end{aligned} \tag{2.52}$$

to derive the steady state

$$\begin{aligned} x_1^* &= \frac{2D(k_u(r_x) p r_x - Dh)}{h(D + h) k_u(r_x)} \\ &= -\frac{2(1 - 2^{-1/n}) \left(D^2 - (2^{1/n} - 1) k_u(r_x) p r_x \right)}{D k_u(r_x)} \end{aligned} \tag{2.53}$$

From which all other (cell) steady states can be calculated using equation 2.45.

The system, in fact, has $n + 1$ steady state solutions: $n - 1$ are complex, one real solution is the trivial extinction steady state, where the cell population gets washed out of the chemostat vessel ($x_i^* = 0$), and the remaining real solution, which we calculated above, is the survival steady state, where the cell population establishes at a non-zero concentration.

The Jacobian matrix for the resident model is

$$\boldsymbol{J} = \begin{bmatrix} -D - h & 0 & 2h & \dots & \frac{D x_1^*}{u^*} \\ h & -D - h & 0 & \dots & \frac{D x_1^*}{2^{1/3} u^*} \\ 0 & h & -D - h & \dots & \frac{D x_1^*}{2^{2/3} u^*} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -k_u(r_x) u^* & -k_u(r_x) u^* & -k_u(r_x) u^* & \dots & -\frac{p}{u^*} \end{bmatrix} \tag{2.54}$$

Evaluating the Jacobian at the survival steady state from the steady state expressions we calculated earlier (after much algebra) shows that the eigenvalue with maximum real part

$$\max_{\lambda \in \Lambda} [Re(\lambda)] = D - \frac{p k_u(r_x) r_x (2^{1/n} - 1)}{D} \tag{2.55}$$

Since a negative real part for the eigenvalue with largest real part implies that all real parts are negative - i.e. the survival condition is stable - we have the survival condition

$$k_u(r_x) r_x (2^{1/n} - 1) > \frac{D^2}{p} \tag{2.56}$$

Note the similarity of this condition with the survival condition from the unstructured model (Equation 2.16). Also note that this condition effectively imposes the same survival requirement of positive initial growth in a chemostat that has reached (nutrient) steady state in the absence of growing cells (see Equation 2.17 in the unstructured model). In other words, the initial overall rate of growth (for the entire population) is

$$k_u(r_x)(r_x)(2^{1/n} - 1) \tag{2.57}$$

which is slower, by a factor of $(2^{1/n} - 1)$, compared to the unstructured population model (Equation 2.17). This growth rate, then, must be greater than the dilution rate, D, if the lineage is to survive in the chemostat.

As with the unstructured system, the structured system always has a single attractor because the extinction condition is simply the reverse of the survival condition. For the remainder of this section we will assume that we are operating in the parameter regime where all r_x support survival.

The eigenvector of the eigenvalue with largest real part is

$$\boldsymbol{x} = \left(\frac{x_1^*}{x_1^*}, \frac{x_2^*}{x_1^*}, \ldots, \frac{x_n^*}{x_1^*} \right) = (1, 2^{\frac{1-2}{n}}, \ldots, 2^{\frac{1-n}{n}}) \tag{2.58}$$

which is the steady state distribution we have calculated earlier.

This means that the cell population approaches the survival steady state in a direction tangent to this slow eigendirection - in other words, the system will quickly converge to the ratios of the invariant steady state distribution in the early phase of the trajectory and then grow along the direction of $\boldsymbol{x}$ until it reaches

the survival steady state (Figure 2.3.2). This is the characteristic behavior of an invariant manifold [116].

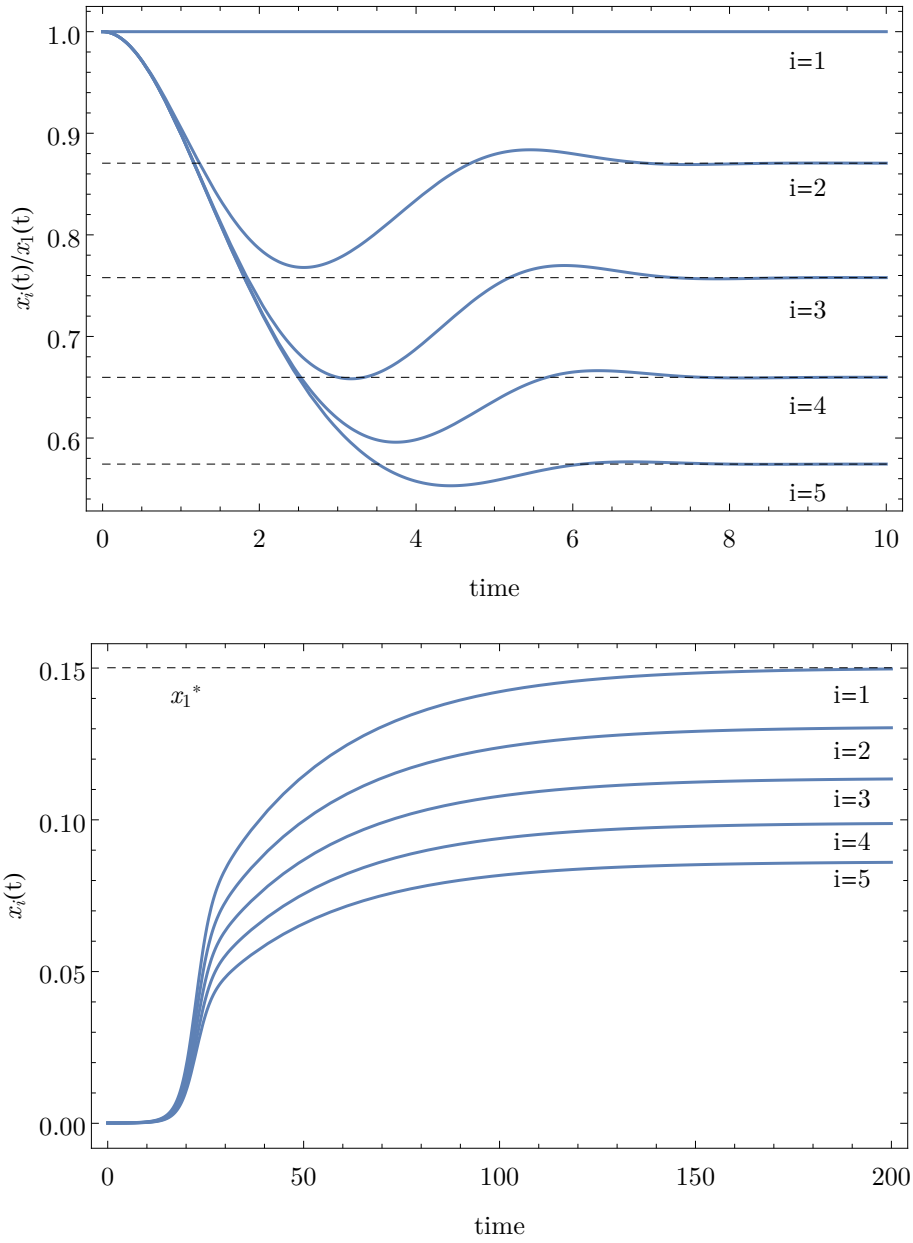


Figure 2.3.2: Simulation of the structured resident model. *Top*: Fast approach to the invariant resident manifold, $\mathcal{M}_R$. Dashed lines show analytical invariant state distribution values (equation 2.58). *Bottom*: Approach to steady state along the invariant manifold. Dashed line shows analytical value of x_1^* from equation 2.53. *Parameters*: $k_u(r_x) = 2\ (ct)^{-1}, r_x = 0.6, p = 0.2\ ct^{-1}, d = 0.03\ t^{-1}, n = 5$. Axes are in arbitrary units of concentration, c, and time, t.

We can take advantage of the invariant property to reduce the structured model on the (resident) invariant manifold, $\mathcal{M}_R$, to the unstructured model, with some modifications - and under the assumption that the invariant manifold always exists, is unique, and globally attractive (not shown). In this case, let

$$\frac{x_i(t)}{x_j(t)} = const. = \frac{x_i^*}{x_j^*} \tag{2.59}$$

for all $t > 0$, where the terms with the asterisk superscript are the invariant ratios calculated earlier. In particular, we have

$$\frac{x_i(t)}{x_1(t)}\Big|_{\mathcal{M}_R} = \frac{1}{2^{(i-1)/n}} \tag{2.60}$$

on the resident manifold, $\mathcal{M}_R$. Then, the structured model can be transformed, after defining $w = 2^{1/n} - 1$, to

$$\begin{aligned}
\dot{x}_1 &= uk_u(r_x)r_x\left[-x_1 + 2x_n\right] - x_1 D \\
&= uk_u(r_x)r_x\left[-x_1 + 2^{1/n}x_1\right] - x_1 D \\
&= uk_u(r_x)r_x x_1\left[-1 + 2^{1/n}\right] - x_1 D \\
&= uk_u(r_x)r_x x_1\left[w\right] - x_1 D \\
\dot{u} &= p - uD - uk_u(r_x)\sum_{i=1}^{n} x_i \\
&= p - uD - uk_u(r_x)x_1\frac{1}{2}\left(1 + \frac{1}{w}\right)
\end{aligned} \tag{2.61}$$

The remaining derivatives are simply multiples of $\dot{x}_1$. As expected, the population size approaches 0 in each cell state as $n \to \infty$, since an increasing amount of nutrient is required to produce a daughter cell.

Solving Equation 2.61 at steady state we obtain the expressions we have calculated earlier (Equations 2.50 and 2.53). The dynamics of the two variable system are identical, except for the initial part of the trajectory while the structured model approaches the invariant manifold (Figure 2.3.3). If the state initial values are chosen to lie on the manifold then the dynamics are identical for $t \geq 0$.

The per-capita growth rate of x_1 in the reduced model is simple (and similar to the one from the unstructured model):

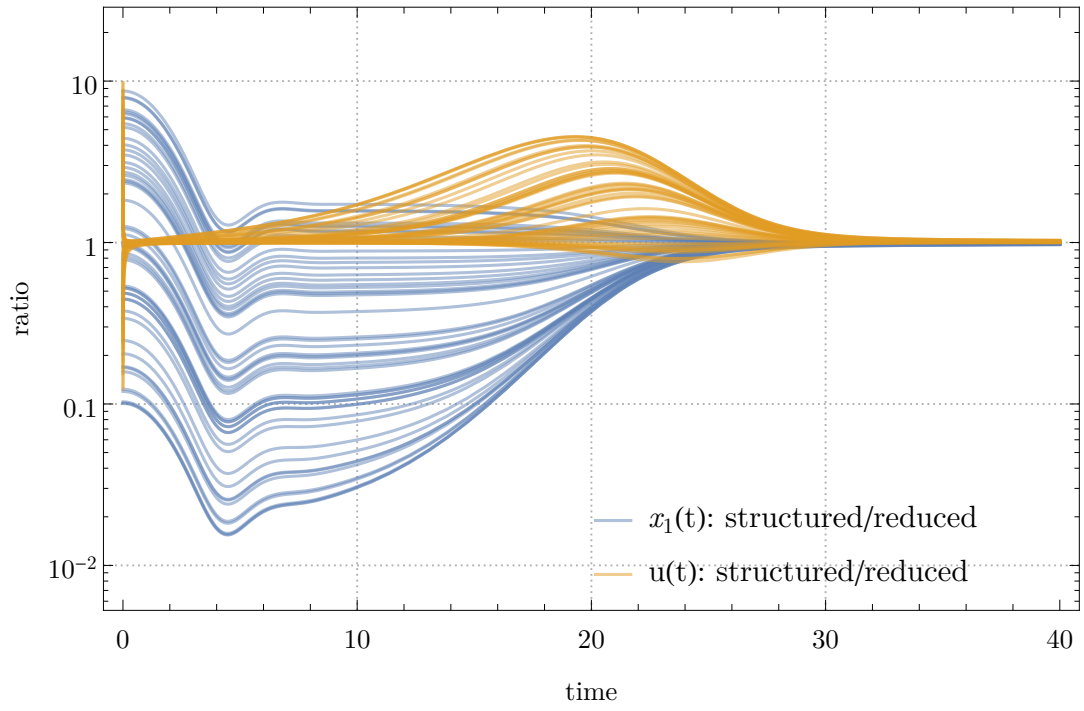


Figure 2.3.3: Comparison of $x_1(t)$ and $u(t)$ between the structured model and the reduced model for 50 random initial conditions. Dynamical trajectories converge soon after the structured model reaches the invariant manifold (approximately $t = 10$, determined by the time taken for all $|\frac{\dot{x}_i}{\dot{x}_j}| < \epsilon$, for a small value of ϵ typically six orders of magnitude smaller than the initial concentration of cells, $x_i(0)$.). The time required to reach steady state in both models is approximately $t = 200$ (determined in a similar manner, but for the $|\dot{x}_i| < \epsilon$). *Parameters*: $k_u(r_x) = 2\ (ct)^{-1}, r_x = 0.6, p = 0.2\ ct^{-1}, d = 0.03t^{-1}, n = 5$.

$$\begin{aligned} \frac{\dot{x}_1}{x_1}\Big|_{\mathcal{M}_R} &= uk_u(r_x)r_x\,[w] - D \\ &= uk_u(r_x)(r_x)\left[2^{1/n} - 1\right] - D \end{aligned} \tag{2.62}$$

The (n) per-capita growth rates of the x_i in the structured model converge to the above expression when the model reaches the invariant state distribution (Figure 2.3.4). This will be useful in later analysis.

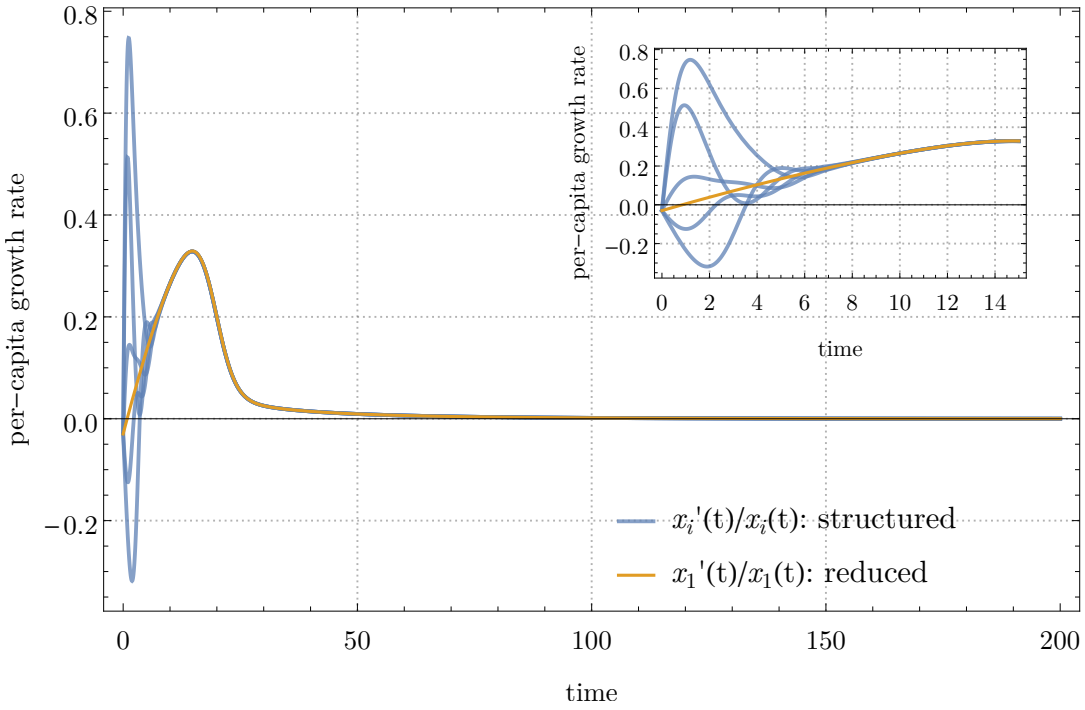


Figure 2.3.4: Comparison of the per-capita growth rate of the reduced model and the per-capita growth rates of the structured model. The per-capita growth rates fluctuate in the early part of the trajectory and converge to the per-capita growth rate of the x_1 in the reduced model when the system reaches the invariant manifold. $t = 200$ is approximately the time required to reach steady state in both models. *Parameters*: $k_u(r_x) = 2\ (ct)^{-1}, r_x = 0.6, p = 0.2ct^{-1}, d = 0.03t^{-1}, n = 5$.

To analyze the outcome of invasion by a rare mutant lineage we will consider the augmented resident-mutant model:

$$\begin{aligned}\dot{\mathbf{x}} &= uk_u(r_x)r_x\mathbf{A}\cdot\mathbf{x} - \mathbf{x}D\\ \dot{u} &= p - uD - u\left(k_u(r_x)\mathbf{1}\cdot\mathbf{x} + k_u(r_y)\mathbf{1}\cdot\mathbf{y}\right)\\ \dot{\mathbf{y}} &= uk_u(r_y)r_y\mathbf{A}\cdot\mathbf{y} - \mathbf{y}D\end{aligned} \tag{2.63}$$

where we have again used y_i as the mutant lineage state concentrations and r_y as the resource allocation parameter of the mutant lineage. To determine whether a nascent mutant lineage will invade, we obtain the model's Jacobian matrix - which is, again, block-triangular with general form:

$$\boldsymbol{J}_{RM} = \begin{bmatrix}\boldsymbol{J}_R & \mathbf{0}\\ \mathbf{0} & \boldsymbol{J}_M\end{bmatrix} \tag{2.64}$$

where the top-left submatrix, $\boldsymbol{J}_R$, is the Jacobian of the resident model (Equation 2.54). Since the mutant lineage invades at the steady state survival point of the resident model (Equation 2.50), the eigenvalues of the $\boldsymbol{J}_R$ submatrix are negative; therefore, we are only interested in the eigenvalues of $\boldsymbol{J}_M$, which has general form:

$$\boldsymbol{J}_M = \begin{bmatrix} -D - r_y k_u(r_y)u & 0 & 0 & \dots & 2r_y k_u(r_y)u \\ r_y k_u(r_y)u & -D - r_y k_u(r_y)u & 0 & \dots & 0 \\ 0 & r_y k_u(r_y)u & -D - r_y k_u(r_y)u & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & -D - r_y k_u(r_y)u \end{bmatrix} \tag{2.65}$$

which is an upper-Hessenberg matrix. Evaluating at the resident model's survival steady state (i.e. $u = \frac{h}{k_u(r_x)r_x}$) and solving the eigensystem we can show that

$$\max_{\lambda \in \Lambda} [Re(\lambda)] = D\left(\frac{k_u(r_y)r_y}{k_u(r_x)r_x} - 1\right) \tag{2.66}$$

is the real part of the eigenvalue with the largest real part. This is the same as the per-capita growth rate of the mutant lineage in the unstructured (Equation 2.21) model. Moreover, it has the same form as per-capita growth rate of the structured model that has been reduced on the (resident) invariant manifold (Equation 2.62).

The eigenvector of this eigenvalue is:

$$\boldsymbol{y} = \left(\frac{y_1^*}{y_1^*}, \frac{y_2^*}{y_1^*}, \dots, \frac{y_n^*}{y_1^*}\right) = (1, 2^{\frac{1}{n}-\frac{2}{n}}, \dots, 2^{\frac{1}{n}-\frac{n}{n}}) \tag{2.67}$$

which is the same as the resident model's survival eigenvector (Equation 2.58). More concretely, the initially scarce mutant lineage first converges to an invariant invasion manifold, $\mathcal{M}_I$, where the per-capita growth rates of the mutant cell states converge to

$$\begin{aligned} \frac{\dot{y}_i}{y_i}\Big|_{\mathcal{M}_I} &= u^{*resident*} k_u(r_y) r_y [D/h] - D \\ &= D\left(\frac{k_u(r_y)r_y}{k_u(r_x)r_x} - 1\right) \end{aligned} \tag{2.68}$$

since the mutant lineage appears in an environment where the nutrient level is fixed (conditioned) by the resident ($u^{*resident*} = \frac{h}{k_u(r_x)r_x}$).

We can interpret the above results as follows: an ancestral (resident) lineage, with phenotype r_x fixes in the chemostat and establishes a steady state nutrient concentration, $u^{*resident*}$. The mutant lineage, with phenotype r_y emerges, and is initially present at very small concentration relative to the ancestral resident. While the mutant lineage is relatively small, it will grow toward an invariant (growth) state distribution - the invariant invasion manifold, $\mathcal{M}_I$. Importantly, the mutant lineage will do so without appreciably changing the amount of nutrient in the environment, which stays fixed at $u^{*resident*}$.

In this case, the invasion manifold, $\mathcal{M}_I$, is equivalent to the mutant manifold, $\mathcal{M}_M$, which is itself equivalent to the resident manifold, $\mathcal{M}_R$. In other words, a lineage's phenotype does not affect the invariant growh state distribution. It is for this reason that the invasion condition is a simple ratio between the on-manifold growth rates of the mutant and resident, i.e.

$$k_u(r_y)r_y > k_u(r_x)r_x \tag{2.69}$$

which can also been written as

$$u^{*resident*} > u^{*mutant*} \tag{2.70}$$

This is the same, nutrient-minimizing, invasion condition as that of the unstructured model (Equation 2.26). Therefore, the adaptive dynamics are qualitative similar between the two models.

In Chapter **??** we will see that the invasion manifold is not necessary equal to the mutant and resident manifolds, particularly for the case where multiple nutrients are present in the environment. In those cases, the invasion manifold exists only transiently and, while it can be used to predict the outcome of invasion, it usually cannot be calculated analytically.

2.4 A structured model with intermediate cellular species

In this section we will analyze a variant of the structured model whereby the process by which cells take up nutrients from the chemostat vessel and metabolize them to cellular resources has been further partitioned to two distinct reaction channels. This modification allows for finer model control and paves the way to models where metabolic tradeoffs can be implemented mechanistically. Because this model is only an intermediate component required for analysis of later models we will keep treatment and exposition as brief as possible.

The reaction channels in this model are shown schematically in Figure 2.4.1. The cell population is now partitioned into the x_i states, which are waiting for nutrients, and the χ_i states, which are 'competent' forms of the cells and that metabolizing nutrients. We again assume that nutrient influx is constant at rate p and both nutrients and cells are diluted at constant rate D. Unlike previous models, however, we will omit the resource-allocation parameter (we will introduce it again in the next section, but in a different incarnation). Also, we have removed the dependence of k_u on r_x, in contrast to the previous model.

$$x_i + u \xrightarrow{k_u} \chi_i \xrightarrow{k_i} x_{i+1}$$

$$x_n + u \xrightarrow{k_u} \chi_n \xrightarrow{k_n} 2x_1$$

$$\emptyset \xrightarrow{p} u$$

$$\{x_i, \chi_i, u\} \xrightarrow{D} \emptyset$$

Figure 2.4.1: Reaction channel schematic for the structured model with intermediate species. There are $2n+1$ variables in the model: n x_i cell states, n χ_i "competent" cell states, and one nutrient variable, u. Note that, unlike the previous model (Section 2.2) the k_u here are not functions of the resource allocation parameter.

The nutrient uptake rate is the same for all x_i states but we have allowed the nutrient metabolism rate, k_i, to be state-specific. The equations for this model are

$$\begin{aligned}
\dot{x}_1 &= -x_1 u k_u + 2\chi_n k_n - x_1 D \\
\dot{x}_i &= -x_i u k_u + \chi_{i-1} k_{i-1} - x_i D, \text{ for } 1 < i \leq n \\
\dot{\chi}_i &= x_i u k_u - \chi_i k_i - \chi_i D \\
\dot{u} &= p - uD - u k_u \sum_{i=1}^{n} x_i
\end{aligned} \tag{2.71}$$

We will primarily be concerned with the steady states of the system, where the following are true

$$\chi_i^* = x_i^* \frac{u^* k_u}{k_i + D} = x_i^* \frac{h}{k_i + D} \tag{2.72}$$

where we have used $h = u^* k_u$, as before. This expression gives the ratio of metabolizing/non-metabolizing cells for class i at steady state. We also have

$$\frac{\chi_{i-1}^*}{x_i^*} = \frac{h + D}{k_{i-1}} \tag{2.73}$$

From which we can derive

$$\frac{x_{i-1}^*}{x_i^*} = \frac{h + D}{h} \frac{k_{i-1} + D}{k_{i-1}} \tag{2.74}$$

Chaining terms we get

$$\frac{x_i^*}{x_n^*} = \left(\frac{h + D}{h} \right)^{n-i} \left[\prod_{j=1}^{n-1} \frac{k_j + D}{k_j} \right] \tag{2.75}$$

The derivatives vanish at steady state, therefore

$$\sum_{i=1}^{n} \dot{x}_i + \dot{\chi}_i = 0 \tag{2.76}$$

Following the methods from Section 2.3, we can expand the summation and rearrange to obtain

$$\begin{aligned} k_n \chi_n^* &= D \sum_{i=1}^{n} x_i^* + \chi_i^* \\ \frac{h}{D} &= \frac{k_n + D}{k_n} \sum_{i=1}^{n} \left(\frac{x_i^*}{x_n^*} + \frac{\chi_i^*}{x_n^*} \right) \end{aligned} \tag{2.77}$$

Substituting the steady state ratios into this expression we get an expression for h:

$$\boxed{h = \frac{D}{(2\alpha)^{1/n} - 1}} \tag{2.78}$$

with

$$\alpha = \prod_{i=1}^{n} \left(\frac{k_i}{k_i + D} \right) \tag{2.79}$$

It is interesting to note the progression of h, which acts a scaling/transformation factor, from the unstructured model ($h = D$, Equation 2.17), to the structured model ($h = \frac{D}{2^{1/n}-1}$, Equation 2.48), to this model (Equation 2.78).

The (survival) steady state solution for the nutrient can then be written as

$$u^* = \frac{D}{k_u \left[(2\alpha)^{1/n} - 1 \right]} \tag{2.80}$$

Moreover, setting

$$\dot{u} = 0 \implies \frac{p - u^* D}{h} = \sum_{i=1}^{n} x_i^* \tag{2.81}$$

Dividing both sides by x_n^* and using the steady state ratios (eq: 2.75) and substituting for h, we can write an expression for the steady state of x_1:

$$x_1^* = \frac{(2\alpha) k_u p - (2\alpha)^{\frac{2}{n}} \left(D^2 + k_u p \right)}{D k_u \sum_{i=1}^{n} \left[(2\alpha)^{\frac{1-i}{n}} \left(\prod_{j=1}^{i-1} \frac{k_j}{D + k_j} \right) \right]} \tag{2.82}$$

from which we can write steady state expressions for all other (cell) variables.

Stability analysis proceeds as with the earlier models so we will forego presenting it here. The important components we require from this intermediate model are the steady state expressions. In particular, we have shown how the steady state of the nutrient scales, relative to the structured model (Equation 2.39), to take into account the rates of nutrient metabolism, k_i, at each step of the growth chain.

2.5 A structured model with alternative metabolic pathways

In this section we will look at a population model where cells metabolize nutrients using two parallel metabolic pathways, each characterized by a set of parameters that describe its rate and yield. The choice of pathway for metabolizing a unit of nutrient is determined by the phenotype of the cell lineage. We will first develop the model and use earlier results to derive expressions that we will ultimately use to analyze the adaptive dynamics of metabolic pathway choice.

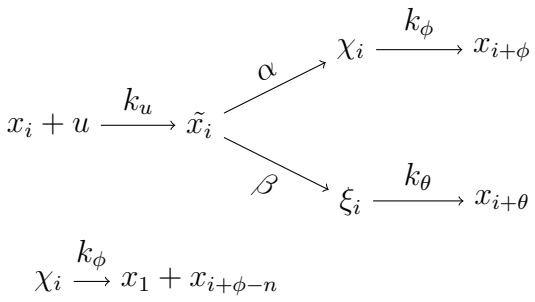
The population is structured into n cell states, as with previous models, but now we will introduce two more intermediate species to represent cells (temporarily) committed to one of the two metabolic pathways. The reaction channels are shown schematically in Figure 2.5.1.

We will call the pathway that transitions from x_i to χ_i the *phi* pathway, and the pathway that transitions from x_i to ξ_i the *theta* pathway. The phi pathway is associated with metabolic rate k_ϕ and yield ϕ, while the theta pathway has rate k_θ and yield θ. Yield is interpreted as the number of states gained for each unit of nutrient metabolized along a specific pathway. When the cells exceed a critical state, n, the mother cell will produce a daughter cell in the first (basal) state and will itself transition to a state whose index is the remainder of the new "state mass" after replication ($-n$). For example, if $n = 10, \phi = 3$ and current state is $i = 9$, metabolism along the phi pathway will result in

$$x_{i=9} \to \chi_{i=9} \to x_1(\text{daughter cell}) + x_2(\text{mother cell})$$

In other words, the daughter cell is always produced in the first state but the mother cell may transition to a state other than the first. This has implications that will become important later on.

The equations for this model are:



$$\chi_i \xrightarrow{k_\phi} x_1 + x_{i+\phi-n}$$

when $i + \phi > n$

$$\xi_i \xrightarrow{k_\theta} x_1 + x_{i+\theta-n}$$

when $i + \theta > n$

$$\emptyset \xrightarrow{p} u$$

$$\{x_i, \chi_i, \xi_i, u\} \xrightarrow{D} \emptyset$$

Figure 2.5.1: Schematic for the structured model with alternative metabolic pathways showing reaction channels. There are two alternative metabolic pathways that we will name the *phi* (ϕ) and the *theta* (θ) pathway. Successful metabolism of one unit of nutrient along the phi pathway forces the cell to transition from state i to state $i + \phi$. We refer to ϕ as the yield of the, similarly named, phi pathway. The same holds for the theta pathway, with yield θ. If the state transition exceeds some critical state, n, then a daughter cell is produced and the mother cell transitions to the state corresponding to the remainder of the "state mass" after replication. There are $3n + 1$ dynamical variables in the model: n x_i "free' cell states; n χ_i cell states committed to the ϕ pathway; n ξ_i cell states committed to the θ pathway; and, one nutrient variable, u. The intermediate species, $\tilde{x}_i$ is at quasi steady state with respect to the rest of the system and acts as a decision point.

$$\begin{aligned}
\dot{\boldsymbol{x}} &= -\boldsymbol{x}(uk_u + D) + k_\phi \boldsymbol{\Phi}\boldsymbol{\chi} + k_\theta \boldsymbol{\Theta}\boldsymbol{\xi} \\
\dot{\boldsymbol{\chi}} &= \boldsymbol{x}(uk_u s_x) - \boldsymbol{\chi}(k_\phi + D) \\
\dot{\boldsymbol{\xi}} &= \boldsymbol{x}(uk_u(1 - s_x)) - \boldsymbol{\xi}(k_\theta + D) \\
\dot{u} &= p - uD - uk_u \mathbf{1} \cdot \boldsymbol{x}
\end{aligned} \tag{2.83}$$

Where we have used the approximation that the $\tilde{x}_i$ intermediates are at quasi steady state with respect to the rest of the system - i.e.

$$\begin{aligned} \dot{\tilde{x}}_i = 0 = x_i u k_u - (\alpha + \beta)\tilde{x}_i \\ \implies \tilde{x}_i = x_i \frac{u k_u}{\alpha + \beta} \end{aligned} \tag{2.84}$$

We have also defined

$$s_x = \frac{\alpha}{\alpha + \beta} \tag{2.85}$$

as the metabolic pathway "preference" parameter. s_x may be interpreted as the success probability of a Bernoulli trial, where success means that the next unit of nutrient is metabolized using the phi pathway and failure means that it is metabolized using the theta pathway. Mechanistically, we can interpret s_x and $1 - s_x$ as a combination of the number and activity of enzymes in a metabolic flux network that divert nutrient metabolism along different pathways [117, 118].

In effect, the $\tilde{x}_i$ intermediates act as decision points in the model, channeling a fraction s_x of the nutrients to the phi pathway and the remainder, $1 - s_x$, to the theta pathway.

The transition matrices, $\mathbf{\Phi}, \mathbf{\Theta}$ have, respectively, elements

$$\boldsymbol{\phi}_{i,j} = \begin{cases} 1 & : j + \phi = i \\ 0 & : \text{otherwise} \end{cases} + \begin{cases} 1 & : j + \phi - n = i \\ 0 & : \text{otherwise} \end{cases} + \begin{cases} 1 & : j + \phi > n \quad \&\& \quad i = 1 \\ 0 & : \text{otherwise} \end{cases} \tag{2.86}$$

$$\boldsymbol{\theta}_{i,j} = \begin{cases} 1 & : j + \theta = i \\ 0 & : \text{otherwise} \end{cases} + \begin{cases} 1 & : j + \theta - n = i \\ 0 & : \text{otherwise} \end{cases} + \begin{cases} 1 & : j + \theta > n \quad \&\& \quad i = 1 \\ 0 & : \text{otherwise} \end{cases} \tag{2.87}$$

for transitions from cell state j to cell state i, and implement the modulo arithmetic of state progression and replication.

A small heuristic model will be defined, before we proceed further, and will be referred to throughout this section as an example, with "structural" parameters: $n = 4, \phi = 1, \theta = 2$. Figure 2.5.2 shows the decision tree that corresponds to

this heuristic system for a (possibly new-born) cell that starts in state 1 and transitions through the state space to replication.

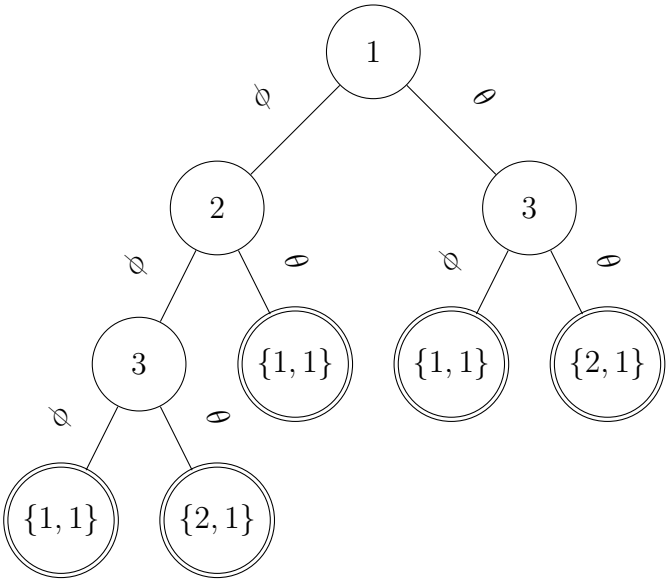


Figure 2.5.2: Decision tree for the small heuristic model for alternative pathways with $n = 3, \phi = 1, \theta = 2$. A new-born mother cell begins in state 1 and progresses through the state space by metabolizing nutrients along the phi or theta pathway. Edge labels show the pathway used to metabolize nutrients. The probability of using the phi pathway is always s_x while the probability of using the theta pathway is always $1 - s_x$, for $s_x \in [0, 1]$. When the cell transitions past state $n = 3$ a daughter cell is produced and the mother keeps the remaining "state mass". The leaves of the tree (double-edged nodes) are the end points of a single progression through the state space and the number of leaves is the number of distinct progression paths through this space. In this example, there are five possible distinct paths.

A path in the decision tree from the root to a leaf is a state progression sequence. Suppose that we knew in advance which path a mother cell would follow, then we could write a smaller system for that path only; for example, for the path $1 \xrightarrow{\phi} 2 \xrightarrow{\theta} 4 \xrightarrow{\phi} \{1, 1\}$ we could write the subsystem with reaction channels:

$$\begin{aligned}
&x_1 + u \xrightarrow{k_u} \tilde{x}_1 \to \chi_1 \xrightarrow{k_\phi} x_2 \\
&x_2 + u \xrightarrow{k_u} \tilde{x}_2 \to \xi_2 \xrightarrow{k_\theta} x_4 \\
&x_4 + u \xrightarrow{k_u} \tilde{x}_4 \to \chi_4 \xrightarrow{k_\phi} 2x_1 \\
&\{x_i, \chi_i, \xi_i\} \xrightarrow{D} \emptyset \\
&\emptyset \xrightarrow{p} u
\end{aligned} \tag{2.88}$$

If we re-number the indices appropriately, and omit the transient species (assuming quasi steady state), then this subsystem is exactly the model of a structured population with intermediates that we analyzed in Section 2.4. For the small heuristic system above, there are five constituent subsystems. These subsystems, however, are not independent: they share at least some cell variables and always share the nutrient variable.

We will now show how we can apply our analysis from Section 2.4 to independent subsystems and then combine results to analyze the model with alternative metabolic pathways under a new restriction: we will require that all paths must terminate at $\{\text{daughter} = 1, \text{mother} = 1\}$. Under the requirement, the start- and end- points of a cell's progression through the cell growth state space are the same, except for a doubling factor, due to division. With this simplifying fact, we can then consider all possible distinct paths through the state space that a new-born cell (in state i) can take to reach the end, with the understanding that all paths share a common start- and end- point. Because these paths, considered in isolation, are identical to the models encountered in the previous section (Section 2.4), we can re-use our results to analyze the individual paths. The task, then, remains to uncover how the paths contribute to the complete model. There are two ways to enforce this requirement (Figure 2.5.3) and the reason for this restriction will be made clear as we work through the method.

To enforce the restriction we can:

1. Force the post-replication state of the mother cell to 1 regardless of how much "state mass" remains after replication.

2. Change the decision tree (by modifying the pathway choice probabilities according to state) so that transitions cannot go beyond state $n+1$. This is not possible for all trees - in fact, this is only guaranteed when $min\{\phi, \theta\} = 1$.

In either case, we proceed by constructing the Q *independent* subsystems (one for each path) of the model. Using the results from Section 2.4, and renaming variables from x_i to z_i and χ_i to w_i; then, the q_{th} subsystem has reaction channels:

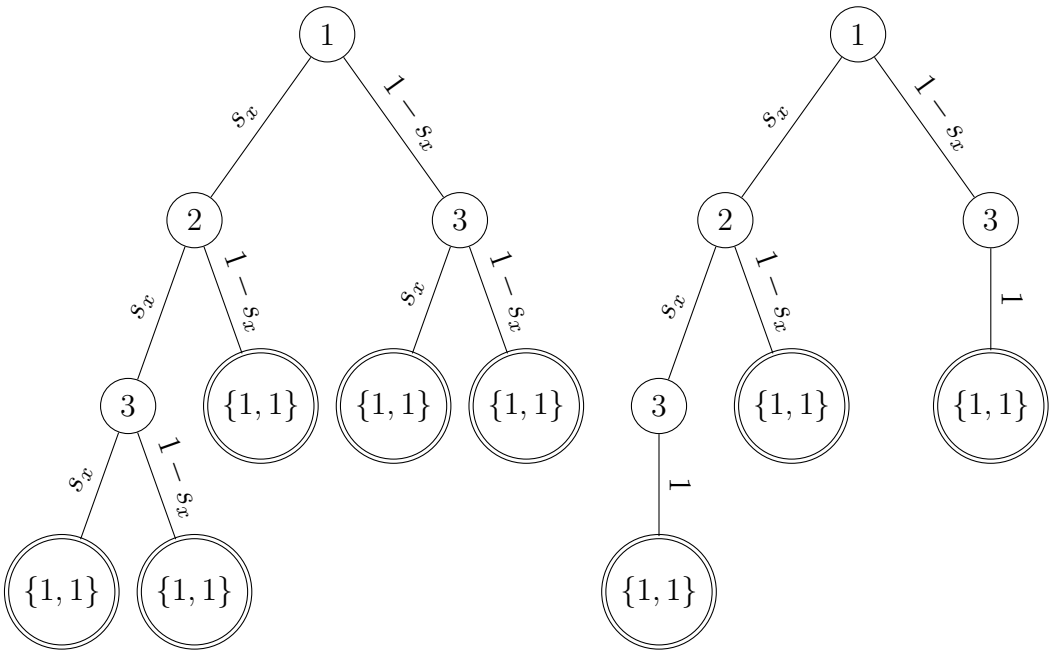


Figure 2.5.3: Decision trees for the small heuristic system with $n = 3, \phi = 1, \theta = 2$ where the leaves are restricted to $\{1, 1\}$. Edge labels are the transition probabilities and the choice of pathway is implied by the magnitude of the state transition. *Left*: Restriction imposed by forcing post-replicative mother cells to transition to state 1. The transition probabilities are not changed. *Right*: Restriction imposed by modifying the decision probabilities so that no transitions past $n+1$ are possible. The probabilities are modified on a per-state basis. Note that there are fewer leaves (paths) in this tree.

$$\begin{aligned} z_{q,1} + u &\xrightarrow{k_u} w_{q,1} \xrightarrow{k_{q,1}} z_{q,2} \\ z_{q,2} + u &\xrightarrow{k_u} w_{q,2} \xrightarrow{k_{q,2}} z_{q,3} \\ &\cdots \\ z_{q,n_q} + u &\xrightarrow{k_u} w_{q,n_q} \xrightarrow{k_{q,n_q}} 2z_{q,1} \\ \{z_i, w_i, u\} &\xrightarrow{D} \emptyset \\ \emptyset &\xrightarrow{p} u \end{aligned} \tag{2.89}$$

The q_{th} subsystem has length n_q (number of states transitions before replication) and has metabolic rates $k_{q,i} \in \{k_\phi, k_\theta\}$ according to the i_{th} decision in the path.

For our small heuristic, for example, we have five subsystems ($Q = 5$) with reaction channels

$$
\begin{aligned}
\text{path 1: } \phi,\phi,\phi &\begin{cases} z_1 + u \xrightarrow{k_u} w_1 \xrightarrow{k_\phi} z_2 \\ z_2 + u \xrightarrow{k_u} w_2 \xrightarrow{k_\phi} z_3 \\ z_3 + u \xrightarrow{k_u} w_3 \xrightarrow{k_\phi} 2z_1 \\ \{z_i, w_i, u\} \xrightarrow{D} \emptyset \\ \emptyset \xrightarrow{p} u \end{cases} \\
\text{path 2: } \phi,\phi,\theta &\begin{cases} z_1 + u \xrightarrow{k_u} w_1 \xrightarrow{k_\phi} z_2 \\ z_2 + u \xrightarrow{k_u} w_2 \xrightarrow{k_\phi} z_3 \\ z_3 + u \xrightarrow{k_u} w_3 \xrightarrow{k_\theta} 2z_1 \\ \{z_i, w_i, u\} \xrightarrow{D} \emptyset \\ \emptyset \xrightarrow{p} u \end{cases} \\
\text{path 3: } \phi,\theta &\begin{cases} z_1 + u \xrightarrow{k_u} w_1 \xrightarrow{k_\phi} z_2 \\ z_2 + u \xrightarrow{k_u} w_2 \xrightarrow{k_\theta} 2z_1 \\ \{z_i, w_i, u\} \xrightarrow{D} \emptyset \\ \emptyset \xrightarrow{p} u \end{cases} \\
\text{path 4: } \theta,\phi &\begin{cases} z_1 + u \xrightarrow{k_u} w_1 \xrightarrow{k_\theta} z_2 \\ z_2 + u \xrightarrow{k_u} w_2 \xrightarrow{k_\phi} 2z_1 \\ \{z_i, w_i, u\} \xrightarrow{D} \emptyset \\ \emptyset \xrightarrow{p} u \end{cases} \\
\text{path 5: } \theta,\theta &\begin{cases} z_1 + u \xrightarrow{k_u} w_1 \xrightarrow{k_\theta} z_2 \\ z_2 + u \xrightarrow{k_u} w_2 \xrightarrow{k_\theta} 2z_1 \\ \{z_i, w_i, u\} \xrightarrow{D} \emptyset \\ \emptyset \xrightarrow{p} u \end{cases}
\end{aligned}
\tag{2.90}
$$

Only paths $1, 3, 4$ are available when the restriction is enforced by modifying the decision tree, whereas all five paths are available when the restriction is enforced by truncating the mother's post-replication state.

Each of the Q subsystems is a model with the same rate equations as 2.71, except for the change of variable names. The q_{th} independent subsystem has nutrient steady state u_q^* according to equation 2.80 using that subsystem's parameters.

Now, define the following:

$$\begin{aligned} f_q(t) &= \frac{d}{u_q(t)} \\ F_q(t) &= f_q(t) + k_u \end{aligned} \tag{2.91}$$

for each of the Q independent subsystems. Moreover, let

$$\begin{aligned} f(t) &= \frac{d}{u(t)} \\ F(t) &= f(t) + k_u \end{aligned} \tag{2.92}$$

be the equivalent quantities for the full, alternative metabolic pathways model. Intuitively, we expect each of the Q independent subsystems to contribute toward an expression for the nutrient concentration of the full model, since they represent the possible paths through the cell state space. Indeed, the subsystems combine in the following manner:

$$1 = \sum_{q=1}^{Q} P_q \left(\frac{F_q(t^\dagger)}{F(t^\dagger)} \right)^{n_q} \tag{2.93}$$

where $t^\dagger$ is the time required for the slowest of the Q independent subsystems to reach steady state. P_q is state progression's probability of the q_{th} subsystem, which is simply the product of all the decision probabilities in the decision tree for the q_{th} leaf. The path probabilities for the small heuristic system are shown in Table 2.2 and Figure 2.5.4 shows nutrient trajectories for example parameters using this system.

Note that $F_q(t^\dagger)$ and $F(t^\dagger)$ are proportional to the invasion fitness, since they contain the reciprocal of the nutrient steady state, of the full model and q_{th} sub-model respectively (plus a constant term). To understand Equation 2.93, it might be helpful to think of fitness as a random variable, $\mathcal{F}$, with support on the F_q. *If* the Q sub-models were independent, then

$$E[\mathcal{F}] = \sum_q P(\mathcal{F} = F_q) F_q = \sum_q P_q F_q$$

would give the expected fitness value for a phenotypic lineage with metabolic pathway preference parameter s_x, which determines the path probabilities, $P_q = P_q(s_x)$.

#	metabolism	"truncated" P_q	"decision" P_q
1	ϕ, ϕ, ϕ	s_x^3	s_x^2
2	ϕ, ϕ, θ	$s_x^2(1-s_x)$	NA
3	ϕ, θ	$s_x(1-s_x)$	$s_x(1-s_x)$
4	θ, ϕ	$(1-s_x)s_x$	$(1-s_x)$
5	θ, θ	$(1-s_x)(1-s_x)$	NA

Table 2.2: Subsystem probabilities for the small heuristic system with $n = 3, \phi = 1, \theta = 2$. There are two sets of path probabilities: one for the "truncated" restriction method, where the mother's state is truncated to state 1 after replication, and one for the "decision" restriction method, where the metabolism probabilities are modified so that no state transition beyond $n+1$ can occur. Because the latter method unbalances the decision tree there are fewer leaves, and therefore fewer state progression paths, compared to the truncation method (see Figure 2.5.3).

However, the sub-models are coupled via the nutrient variable and the start- and end- points in state space; therefore, we should not expect that the above equation must hold. However, Equation 2.93, suggests the form of an 'effective' fitness for each q, which takes into account the coupling of the sub-models. The expectation of the effective fitness, corrected for coupling, gives the fitness of the full model:

$$E[\mathcal{F}_{\text{effective}}] = \sum_q P_q F_q^{n_q} \cdot \frac{1}{F^{n_q-1}} = F \tag{2.94}$$

which we can obtain from Equation 2.93 by multiplying both sides by F. When all paths have the same length, $n_q = n$, the above expression simplifies to

$$F = \left(\sum_q P_q F_q^n \right)^{1/n}$$

It was not possible to obtain a simpler expression for $F(t)$ in general, however, since the solution is the root of a, possibly large, polynomial. Nevertheless, our attempt provides an alternative way to solve for the steady state of the full model by solving a single polynomial involving terms that we can easily calculate. It is straightforward to numerically solve the equation for $F(t^\dagger)$ since we have simple expressions for the $u_q(t^\dagger)$, according to equation 2.80, and therefore for the $F_q(t^\dagger)$.

In particular, we can numerically construct the function

$$u^* = U(s_x) = U(u_1^*, \ldots, u_Q^*) = U(\mathcal{U}_1(s_x), \ldots, \mathcal{U}_Q(s_x)) \tag{2.95}$$

which gives the nutrient steady state, as a function of the metabolic pathway preference parameter, s_x, from the single real root of Equation 2.93. Here we have used

$$\mathcal{U}_q(s_x) = P_q F_q^{n_q} = P_q(s_x)\left(\frac{d}{u_q^*} + k_u\right) \tag{2.96}$$

and we can use the analytical expressions for the q_{th} independent subsystem (equation 2.80):

$$u_q^* = \frac{D}{k_u\left[(2\alpha_q)^{1/n} - 1\right]}$$

Figure 2.5.4 compares simulations of the alternative metabolic pathways model (orange line) with combined simulations of its Q constituent subsystems (blue line, equation 2.96) showing step-like convergence of the two solutions. This is, again, suggestive that a correction to Equation 2.93, or perhaps initial conditions chosen to lie on an invariant manifold, might allow convergence for all t.

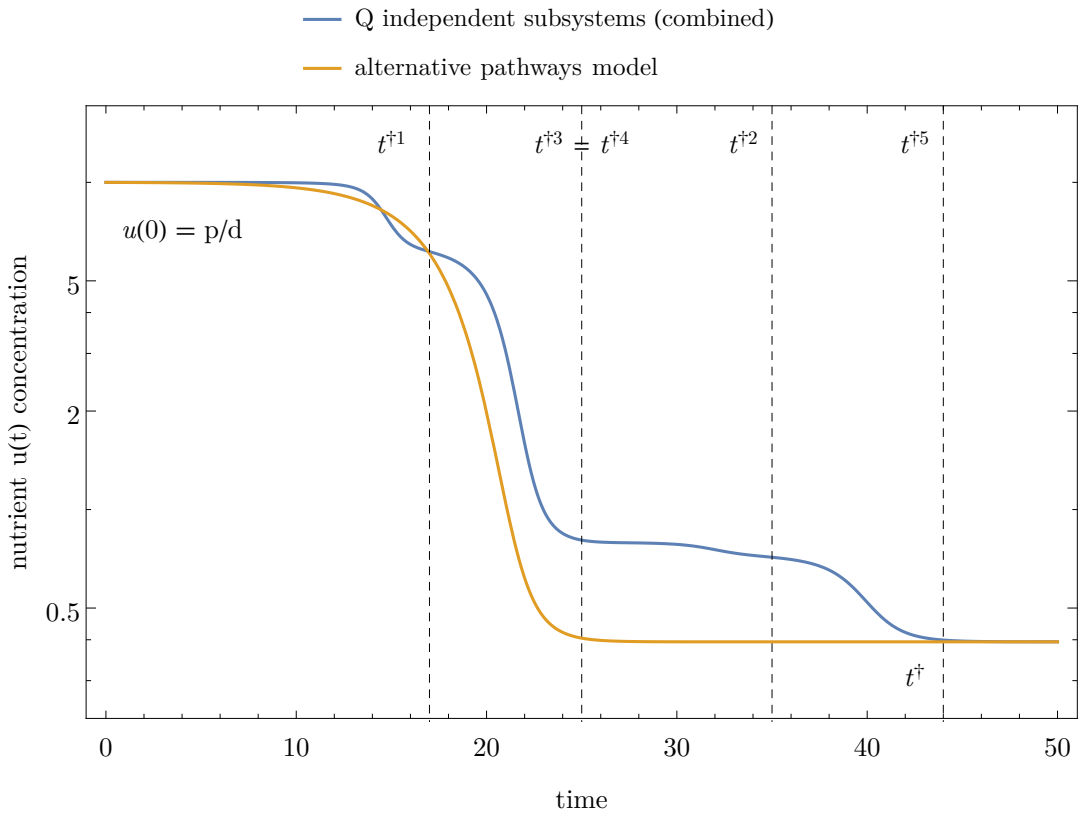


Figure 2.5.4: Simulation of the full alternative metabolic pathways model (equations 2.83 compared with simulations of $Q = 5$ independent subsystems combined according to Equation 2.93, for the small heuristic system with $n = 4, \phi = 1, \theta = 2$. Post-replicative mother cell states are truncated to the first state (Figure 2.5.3, left panel). Equation 2.93 was solved for: $k_u = 0.843\ (ct)^{-1}, k_\phi = 10\ t^{-1}, k_\theta = 1\ t^{-1}, p = 1ct^{-1}, d = 0.1t^{-1}$ and an expression for $u(t)$ was calculated in terms of $u_q(t)$, the nutrient trajectories of the Q subsystems. The $u_q(t)$ were obtained by numerically integrating the independent subsystems. Both the alternative pathways model (blue line) and the combined subsystems (orange line) converge to the analytical value of u^* for $t \geq t^\dagger$. The dashed vertical lines indicate approximate times when each of the $Q = 5$ subsystems, numbered according to Table 2.2, reach their own steady state. When the slowest subsystem reaches its steady state ($t^{\dagger 5}$) the two simulations converge. This suggests that a modification of Equation 2.93 might also correct for the dynamics of the nutrient trajectory.

Because the reciprocal of the nutrient steady state provides a simple measure of invasion fitness (Sections 2.2, 2.3, 2.4) a deconstruction of the progression of a mother cell through state space allows us to quantify the fitness contribution of each of the possible paths to the overall invasion fitness for a particular phenotype value (s_x).

The full model of alternative metabolic pathways permits the cell to make a "decision" about how to metabolize a unit of nutrient according to the outcome of a Bernoulli trial with success probability s_x. This is the simplest, in some sense, model for two alternative options we could have designed. It is a naive model because, unlike real cells, a cell in our model cannot regulate the decision process - except at evolutionary scales, which we will look at in the next subsection. In Figure 2.5.5 the full model's nutrient steady state is shown with a thick black line using our small example system. The nutrient steady state is minimized at approximately $s_x = 0.25$; in other words, when one-quarter of the nutrients, on average, are metabolized using the phi pathway. This is also the phenotype value at which invasion fitness is maximized (see next subsection). However, we can also see that, were the cell capable of self-regulation, it would have done better than any value of s_x by alternating between pathways for incoming nutrients (green line, paths ϕ, θ and θ, ϕ) - see Table 2.3.

path #	u_q^*	$P_q(s_x = 0.280)$
path 1: ϕ, ϕ, ϕ	0.479	0.020
path 2: ϕ, ϕ, θ	0.558	0.0564
path 3: ϕ, θ	0.347	0.202
path 4: θ, ϕ	0.347	0.202
path 5: θ, θ	0.415	0.519
full model ($s_x = 0.280$)	0.394	

Table 2.3: Nutrient steady states for the full alternative metabolic pathways model and its $Q = 5$ constituent subsystems, for the small heuristic system. The nutrient steady state for the full model is evaluated at its optimum (minimum) value, which occurs when $s_x = 0.280$.

The implication of this result is that, even in a genetically clonal (identical) population of cells, all characterized by preference parameter s_x, we should expect some heterogeneity in terms of individual-based fitness: cells that, by chance, have been alternating their pathway choice will have a higher fitness than the

average population and will produce more daughter cells. We can also interpret this result as strong evidence for the selective advantage of metabolic regulation - which is prevalent in biological systems - over naive choice [119, 120, 121, 122].

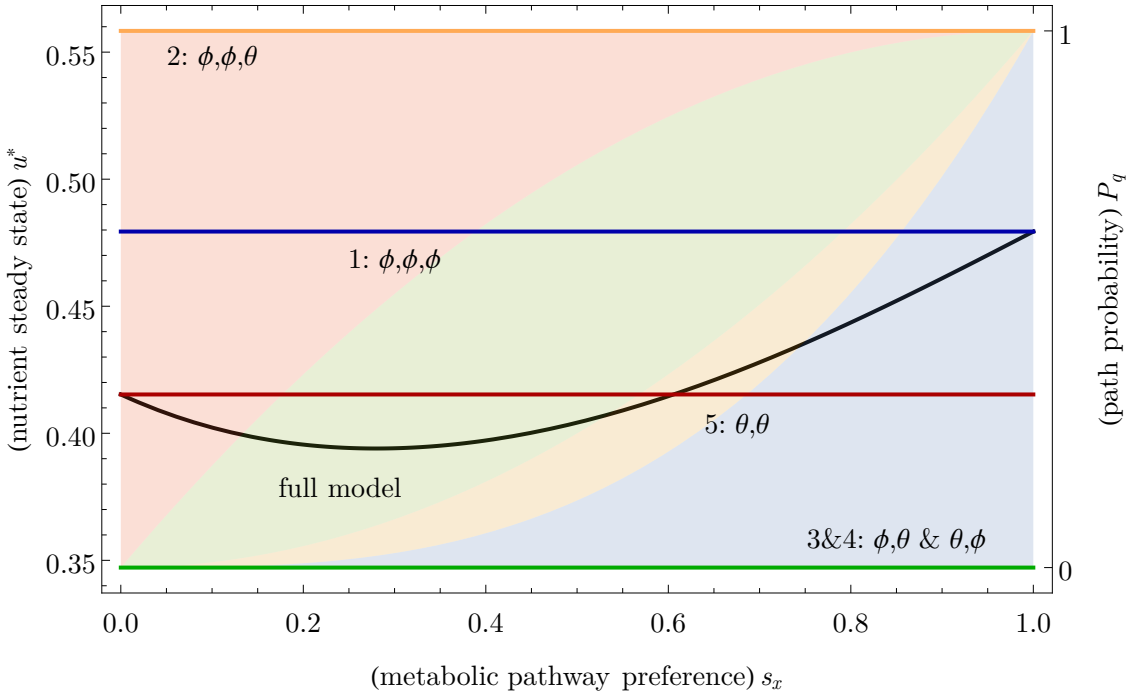


Figure 2.5.5: Comparison of nutrient steady state, u^*, between the full alternative metabolic pathways model, with pathway preference parameter s_x, and its Q independent subsystems. *Parameters:* $n = 3, \phi = 1, \theta = 2, k_u = 0.843\ (ct)^{-1}, k_\phi = 10\ t^{-1}, k_\theta = 1\ t^{-1}, p = 1ct^{-1}, d = 0.1t^{-1}$. The mother post-replicative state was truncated to the first state (Figure 2.5.3, left), which implies that there are $Q = 5$ state progression paths (Table 2.2). The thick black line shows the nutrient steady state as a function of the metabolic preference parameter, s_x, which is the probability that a unit of nutrient is metabolized using the phi pathway; equivalently, $1 - s_x$ is the probability that the nutrient is metabolized using the theta pathway. Solid horizontal lines are the nutrient steady states of the five independent subsystems, which are labelled according to the metabolic processes (note that the third and fourth path have the same steady state). While the nutrient steady states of these subsystems do not change with s_x, the probability that a particular state progression is realized does change with s_x (background colors). The filled curves in the background show the path probabilities, and are color-coded to match the horizontal lines - *blue*: path 1 (ϕ, ϕ, ϕ), *orange*: path 2 (ϕ, ϕ, θ), *green*: paths 3 and 4 $(\phi, \theta\ \&\ \theta, \phi)$, *red*: path 5 (θ, θ).

Similar results are obtained for the alternative metabolic pathways model where

post-replicative mother cells are restricted to the first state by modifying the decision probabilities (not shown).

It was not possible to find a way to combine models of independent subsystems (paths) for the full, unrestricted, alternative pathways model (Figure 2.5.2), because of the 'aperiodicity' of the possible paths. Without the restriction that all cells (re-)enter the state space at the first state we would have to expand the decision tree until all leafs terminate at a common state, possibly accounting for several replication events; however, this is not possible because a cell may never enter the branch(es) that terminate at the desired state - resulting in possibly infinite trees. Nevertheless, because the trees are recurrent - the subgraph under any node is identical to the subgraph of the same node in another position in the tree (see Figure 2.5.2) - it is tempting to speculate that a recurrence-based method to combine subsystems constructed from independent paths in the full model does exist.

2.5.1 Adaptive dynamics in the evolution of metabolic pathway preference

The resident-mutant model with alternative metabolic pathways is the augmented form of equations 2.83:

$$\begin{aligned}
\dot{\boldsymbol{x}} &= -\boldsymbol{x}(uk_u + D) + k_\phi \boldsymbol{\Phi\chi} + k_\theta \boldsymbol{\Theta\xi} \\
\dot{\boldsymbol{\chi}} &= \boldsymbol{x}(uk_u s_x) - \boldsymbol{\chi}(k_\phi + D) \\
\dot{\boldsymbol{\xi}} &= \boldsymbol{x}(uk_u(1 - s_x)) - \boldsymbol{\xi}(k_\theta + D) \\
\dot{u} &= p - uD - uk_u \mathbf{1} \cdot \boldsymbol{x} \\
\dot{\boldsymbol{y}} &= -\boldsymbol{y}(uk_u + D) + k_\phi \boldsymbol{\Phi\zeta} + k_\theta \boldsymbol{\Theta\omega} \\
\dot{\boldsymbol{\zeta}} &= \boldsymbol{x}(uk_u s_y) - \boldsymbol{\zeta}(k_\phi + D) \\
\dot{\boldsymbol{\omega}} &= \boldsymbol{x}(uk_u(1 - s_y)) - \boldsymbol{\omega}(k_\theta + D)
\end{aligned} \tag{2.97}$$

where we have used $\boldsymbol{y}, \boldsymbol{\zeta}, \boldsymbol{\omega}$ for the mutant cell variables and s_y for the mutant metabolic pathway preference parameter.

Finding an exact value for the maximum real part of this model's Jacobian is difficult because of the size of the model. Numerical simulations, however, suggest that

$$Sign[\max_{\lambda\in\Lambda}\{Re[\lambda]\}] = Sign[u_x^* - u_y^*] \tag{2.98}$$

which implies, not unexpectedly, that the invasion condition is again

$$u_x^* > u_y^* \tag{2.99}$$

where u_x^* and u_y^* are interpreted as the nutrient steady states of the resident-only and mutant-only (i.e. the resident model with $s_x \to s_y$) models respectively. We know how to calculate the nutrient steady states for the resident models from the previous subsection, so we are ready to look at the evolutionary dynamics of pathway preference under different parameter regimes.

Figure 2.5.6 shows that the evolutionary optimality of pathway preference between a low-rate, high-yield pathway and a high-rate, low-yield pathway depends on the dilution rate of the chemostat vessel. Environments where the cells are faced with a low dilution rate drive the metabolic pathway preference toward slower metabolisms with higher yields; conversely, when the dilution rate is high, lineages with phenotypes that rely on faster metabolisms with lower yields will be selected.

Similar theoretical results have been obtained using minimal mechanistic models of self-replicating systems that are supported by experimental evidence [109]. The selection pressures, in this case, imply that the optimal strategy may be inefficient with regard to maximizing biomass per unit nutrient; indeed, in the case of the highest dilution rates, the evolutionary optimal metabolic strategy is one of the *least* efficient in converting nutrient to biomass (Figure 2.5.6 inset).

Nevertheless, the dependence of the evolutionarily singular strategy (ESS) on model parameters, particular metabolic pathway parameters is not trivial. Random sampling of the parameter space (Figure 2.5.7) indicates that the majority of parameter sets select for extreme ESS values ($s_x = 0$ or $s_x = 1$), suggesting that metabolic pathway specialists are, in general, likely to be favored across a random assortment of environments and pathways. About 2% of parameter sets had ESS values that were not extremes, favoring mixed metabolic strategies; in particular, parameter sets where the metabolic pathways fell into a "low-yield, high-rate" versus "high-yield, low-rate" combination, were more likely to have intermediate evolutionary optima. Indeed, the example from Figure 2.5.6 was

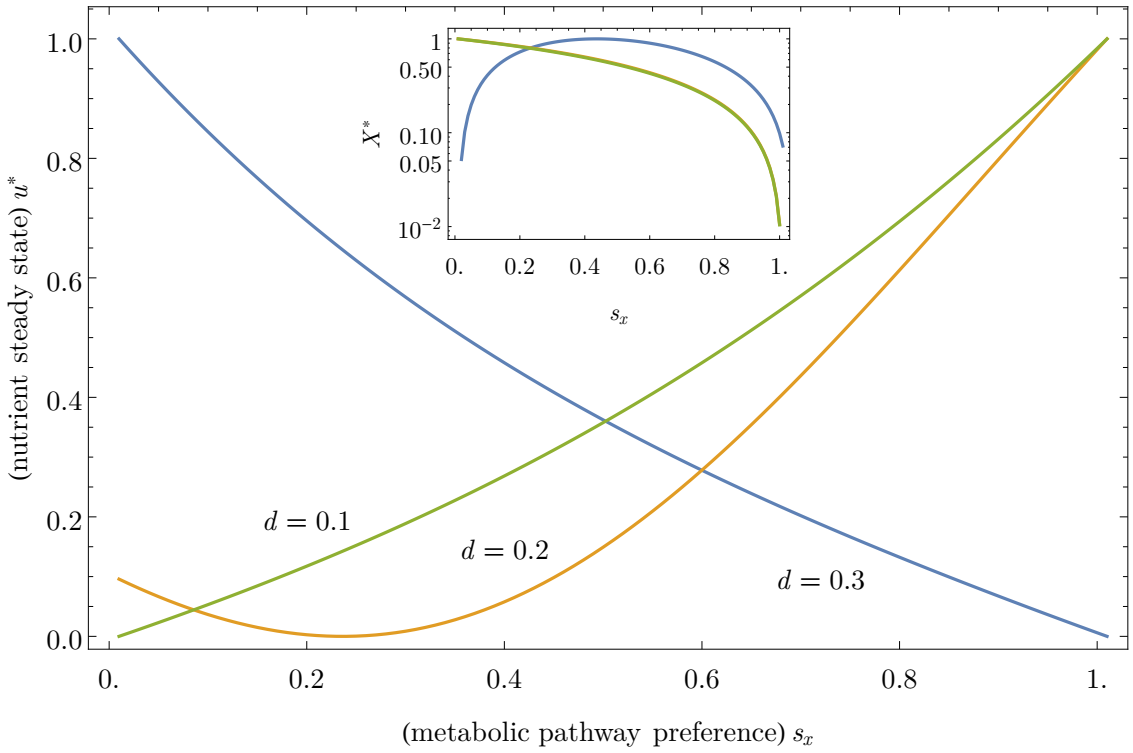


Figure 2.5.6: Evolutionary optimality of pathway preference between a low-rate, high-yield pathway (phi, $s_x \to 0$) and a high-rate, low-yield pathway (theta, $s_x \to 1$) for three different dilution rates, d (line labels). Phenotype values with lower nutrient steady states (u^*) can invade lineages with higher nutrient steady states; therefore, the value of s_x where $u^*(s_x)$ is minimized is a global evolutionary optimum. Environments with high dilution rates select for cells that favor high-rate, low-yield pathways; conversely, low dilution rates will select for cells that favor low-rate, high-yield pathways. Note that the nutrient steady state concentrations have been individually rescaled (linearly) to lie in the $[0, 1]$ interval. *Inset*: The steady state population size (X^*) as function of the s_x phenotype for the three dilution rates, expressed as a fraction of the maximum size for each condition. The population size is at steady state is proportional to the efficiency of converting nutrient to biomass. The evolutionarily optimal strategy does not necessarily coincide with the optimal biomass production strategy.

chosen from among these cases to illustrate the ESS variability in this parameter region. When the metabolic pathways themselves face a tradeoff between rate and yield (i.e. when a pathway has high rate *or* high yield), cells that employ a mixed metabolic strategy are more likely to be evolutionarily successful.

The log ratio of the pathway yields (ϕ/θ), together with the log ratio of the pathway rates (k_ϕ/k_θ), and, to a lesser extent, the dilution rate (D), determined the evolutionary optimality of metabolic strategies, while the nutrient uptake rate

(k_u) and nutrient influx rate (p) did not have a large effect (Figure 2.5.8). We could not find a better parameter transformation to more correctly predict the ESS and separate the two dominant solutions. The difficulty of obtaining an analytical solution for the invasion fitness suggests that a closed-form transformation might not exist.

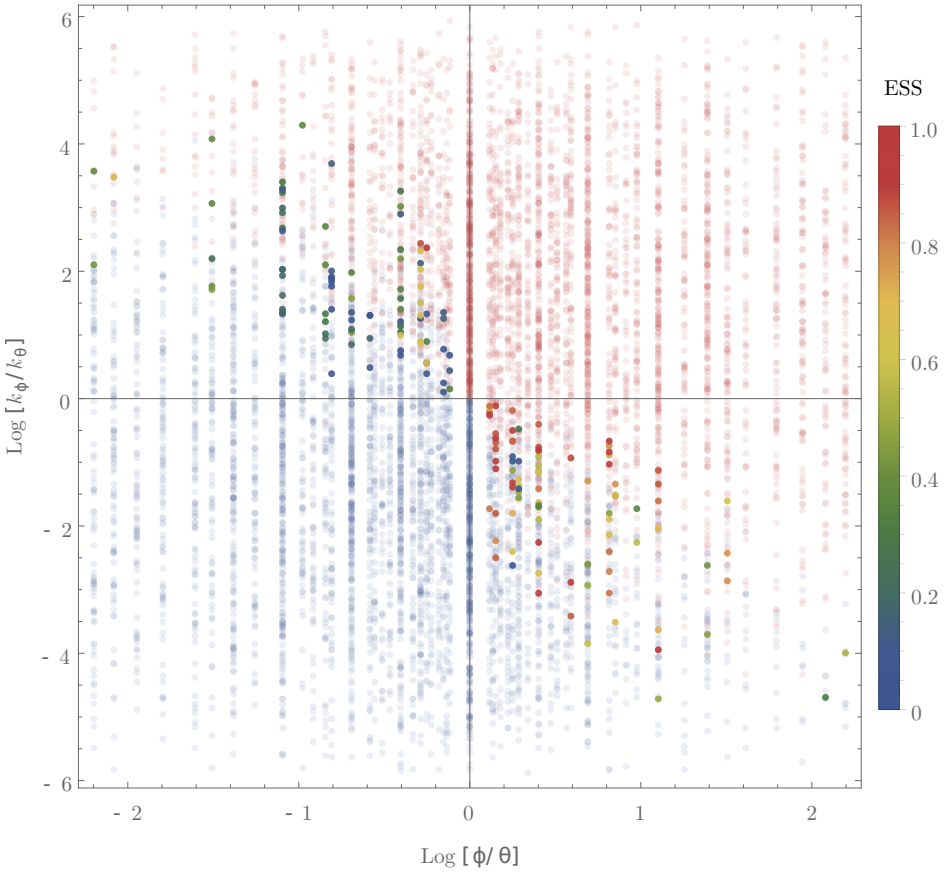


Figure 2.5.7: Random parameter sampling and ESS (global nutrient steady state minimum) determination for the alternative metabolic pathways model. Horizontal axis is the log yield ratio between the two pathways, and vertical axis the log rate ratio. Point color indicates the ESS. The majority of parameter sets have extreme evolutionary optima, $s_x = 0$ (dark blue, exclusively use theta pathway) or $s_x = 1$ (dark red, exclusively use phi pathway). The interface between the pathway "specialists" contains a region of ESS variability - whereby "generalists" may be favored by natural selection (accented points). The interface corresponds to a region where the metabolic pathways lie on a tradeoff curve between yield and rate - in other words, pathways provide, alternatingly, either a high metabolic yield *or* high metabolic rate, but not both. *Method*: $10,000$ parameter sets were randomly sampled. The corresponding models were generated and the global minimum of the nutrient steady state was numerically determined. Parameter sets that did not support growth were rejected.

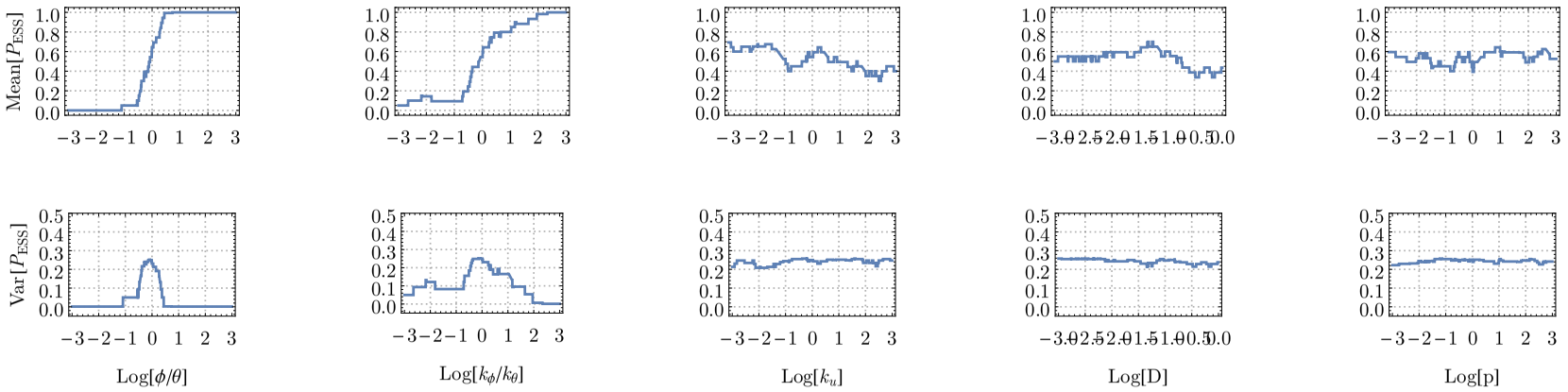


Figure 2.5.8: Mean (top row) and variance (bottom row) of the marginal probability distribution for the evolutionarily singular strategy (ESS) using five log-transformed features according to a k nearest-neighbors model trained on the random parameter sampling dataset. The features are (from left to right): yield ratio, rate ratio, nutrient uptake rate, dilution rate, influx rate. The yield and rate ratios strongly, and independently, influence the ESS; moreover, for these two features, variance is maximized when the log-ratios are close to 0, which indicates that, in these regions, a single feature alone cannot accurately predict the ESS. The ESS ranges between $s_x = 0$ (theta pathway specialist) and $s_x = 1$ (phi pathway specialist).

2.6 Summary of adaptive dynamics models

In this chapter we extended the standard chemostat model of microbial growth to incorporate phenotype-determined cellular resource allocation tradeoffs. We applied the theory of adaptive dynamics to incrementally more complex models, culminating in a theory of microbial metabolic pathway choice.

In Section 2.2 we derived the selection derivative and invasion condition for a simple, unstructured, resource-consumer model where the cell allocates a fraction of its resources toward biomass synthesis and the remaining resources toward driving the nutrient-use pathways. A trade-off between these two cellular processes has been previously experimentally demonstrated [97,98,99]. The invasion condition was in accordance with the R^* concept from the graphical theory of resource competition [42]. We also qualitatively mapped the global adaptation dynamics for an example parameter set, where we made the observation that the long-term outcome of adaptation must dependent on the statistics of the mutational process and the (initial) value of the ancestral phenotype. We will further explore these ideas numerically in Chapters **??** and **??**

We imposed a structure on the population of cells in Section 2.3 and, following the methods from Section 2.2, we found that:

1. The model converges to an invariant manifold, which contains the model's steady state.

2. The structured model can be reduced to the unstructured model on the invariant manifold, through a transformation that accounts for the length of the growth chain.

3. A nascent (scarce) mutant lineage will invade along an invariant invasion manifold, which, for this model is equivalent to the resident model's invariant manifold. We will see in Chapter **??**) that this equivalence does not always hold.

4. As a consequence of point 3, the invasion condition is the same as that of the unstructured model; therefore, the adaptive dynamics are also similar.

In Section 2.4 we included intermediate cell species to more finely describe the

conversion of nutrients to biomass along the growth chain. We derived a second transformation that maps this model to the unstructured model from Section 2.2.

In the final Section (2.5) we constructed a structured-population, resource-consumer, model whereby the nutrients are metabolized using one of two metabolic pathways. In this model, a lineage's phenotype determined the distribution of nutrients between the two pathways.

We showed that, under certain assumptions, the overall fitness of phenotype can be understood in terms of the possible combinations in which nutrients can be metabolized. In other words, the overall fitness of a population of cells with a certain phenotype can be described in an ergodic manner from the fitness of individual cells. We also demonstrate one way in which population heterogeneity may arise from the random metabolic paths that cells realize through the growth chain. This also provides implicit evidence for the selective advantage of metabolic regulation, which would favor the cells that realize the metabolic paths with higher fitness.

In Section 2.5.1 we present theoretical evidence that agree with experimental observations that microbes cultured in slow dilution environments (i.e. low growth rate) employ a low-rate/high-yield metabolic strategy, whereas fast dilution environments (i.e. high growth rate) prefer high-rate/low-yield metabolisms [109].

Finally, we conducted a numerical exploration of the parameter space where we calculated globally adaptively stable metabolic strategies. We found that metabolic specialists dominate the parameter space. A notable exception, where generalist strategies are adaptively stable, occurs when the neither of the two metabolic pathways can maximize both metabolic yield and rate. This limited numerical exploration will serve as a prelude to the more comprehensive numerical methods of Chapters **??** and **??**.

Chapter 3

Modeling and simulation software for adaptive dynamics

3.1 Introduction

In this chapter we will describe the numerical methods developed for generating, simulating, and analyzing models of adaptation dynamics of the type encountered in Chapter **??**. Our aim was to provide a set of tools for working with analytically intractable, but more complex, models (see applications in Chapter **??**). The software was written in the Wolfram Language with an interface provided through Mathematica [123].

The numerical methods were created with the following objectives in mind:

1. Using a flexible specification of a generalized model of adaptation (Section 3.2), the software should be able to automatically generate the dynamical models required to characterize the adaptation process (Section 3.3).

2. Following the framework of adaptive dynamics [69, 73, 74], the adaptation process can be understood in terms of 'invasion events'. Methods must be provided for generating the event's corresponding model (objective 1), simulating (Section 3.3.1), and resolving (Sections 3.3.2 and 3.3.3) the outcome of the invasion simulation.

3. Each step in the treatment of an invasion event (objective 2) should be transparent and verifiable to allow comparisons with theory. Accordingly,

the software must provide a 'supervised' work flow that permits the user to change the parameters and options of the numerical methods on a case-by-case basis (Section 3.3).

4. On the other hand, obtaining a comprehensive description of the adaptation process for a model requires the generation, simulation, and inspection of a large number of models. After an initial round of formulating a model, testing, and possibly comparing with existing theoretical results, the software should provide 'unsupervised' methods to automate the work flow (Section 3.4). These methods implement various approaches for sampling (Section 3.4.1) and systematically mapping (Section 3.4.2) the model's adaptation space.

5. Given an ensemble of adaptation samples, or a complete map (objective 4), we require methods to interpret the results. In our software, we have decided to use graphs to visualize the adaptation process (Section 3.4.3) and discrete Markov processes (Section 3.4.4) to analyze the outcome and dynamics of adaptation.

3.2 Adaptive dynamics models

We consider dynamical non-linear models of microbial population dynamics, similar to those encountered in Chapter **??**, with general form

$$\begin{aligned} \frac{d\mathcal{X}}{dt} &= \boldsymbol{F}_{\mathcal{X}}(\mathcal{X}, \boldsymbol{\nu}, t; \boldsymbol{m}, \mathcal{A}) \\ \frac{d\boldsymbol{\nu}}{dt} &= \boldsymbol{F}_{\boldsymbol{\nu}}(\mathcal{X}, \boldsymbol{\nu}, t; \boldsymbol{m}, \mathcal{A}) \end{aligned} \tag{3.1}$$

The model variables have been partitioned into the set of lineage variables, $\mathcal{X}$, and the set of environment variables, $\boldsymbol{\nu}$; likewise, the parameters have been partitioned into the set of model parameters, $\boldsymbol{m}$, and the set of adaptation parameters, $\mathcal{A}$.

Each lineage in the model, $\ell_i \in \mathcal{L}$, is associated with

1. A vector of lineage variables, $\boldsymbol{x}_i(t) \in \mathcal{X}$.
2. A tuple of adaptation parameter vectors, $\boldsymbol{\alpha}_i \in \mathcal{A}$.

Each of the adaptation parameter vectors, $\boldsymbol{a}_{i,j}$, in the tuple $\boldsymbol{\alpha}_i$ lies on the unit simplex, S. That is, we have

$$\{a_{i,j,1}, a_{i,j,2}, \dots\} \in \boldsymbol{a}_{i,j}$$

such that

$$a_{i,j,1} + a_{i,j,2} + \cdots = 1 \text{ and } a_{i,j,k} \leq 1 \; \forall \; k$$

The position of a lineage in adaptation space, $\mathcal{S}$, therefore, is given by the union of its position in all the unit simplexes, S_K; moreover, $\boldsymbol{\alpha}_i$ is unique to ℓ_i, and lineage identity is indexed by its position in adaptation space.

The grouping of adaptation parameter permits the implementation of trade-offs by explicitly restricting some groups of parameters to lie on the simplex surface. For example, suppose we are modeling the adaptation dynamics for the partitioning of the cell surface between nutrient transporters (phenotype trait A) and the partitioning of ribosome activity between classes of proteins (phenotype trait B). Suppose, further, that there are only two types of nutrients and three classes

of proteins; therefore, two pairs of adaptation parameter vectors - one for each phenotype - are required. Because the cell surface area and the number of ribosomes are finite and independent resources to be partitioned between competing processes we have

$$\begin{aligned} \{a_1, a_2\} \in S_A \quad \text{s.t} \quad a_1 + a_2 = 1 \text{ and } 0 \leq a_1, a_2 \leq 1 \\ \{b_1, b_2, b_3\} \in S_B \quad \text{s.t} \quad b_1 + b_2 + b_3 = 1 \text{ and } 0 \leq b_1, b_2, b_3 \leq 1 \end{aligned} \tag{3.2}$$

Where S_A is the two-dimensional unit simplex and S_B is the three-dimensional unit simplex.

To make the connection to the models we have encountered so far, the growth state-structured model of resource allocation (Equation **??**) can be cast in the required form by setting:

$$\begin{aligned} \mathcal{X} &= \{(x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n), \ldots\} \\ &= \{(x_{\ell,1}, x_{\ell,2}, \ldots, x_{\ell,n})\}, \text{ for each resident/mutant lineage } \ell \in \mathcal{L} \end{aligned}$$

as the lineage variables. There is a single environmental variable – the nutrient – so we have

$$\nu = u$$

The forms of the derivative functions are taken directly from Equation **??** and are:

$$\begin{aligned} F_{x_{\ell,1}} &= u k_u(r_\ell) r_\ell(-x_{\ell,1} + 2x_{\ell,n}) - x_{\ell,1} D \\ F_{x_{\ell,i}} &= u k_u(r_\ell) r_\ell(-x_{\ell,i} + x_{\ell,i-1}) - x_{\ell,i} D, \text{ for } 1 < i \leq n \\ F_u &= p - uD - u \sum_{\ell \in \mathcal{L}} k_u(r_\ell) \sum_{i=1}^{n} x_{\ell,i} \end{aligned}$$

for the ℓth resident/mutant lineage. The model's parameters are

$$\boldsymbol{m} = \{k_u(r_\ell), p, D\}, \text{ for } \ell \in \mathcal{L}$$

Finally, there is a single evolvable parameter simplex, for the trade-off between biomass synthesis and nutrient uptake, so we can write:

$$\mathcal{A} = \{(r_x, 1 - r_x), (r_y, 1 - r_y), \dots\} = \{(r_\ell, 1 - r_\ell)\}\,, \text{ for } \ell \in \mathcal{L}$$

To complete the model definition any method or combination of methods can be used to construct the model equations (matrix, rule-based, etc. . .). The only restriction, however, is that it must be made clear to the software how lineages may be added to and removed from the model; usually, it is enough to include a `$lineage` index variable in the equations and a `$number_lineages` in the model parameters. Accordingly, when `$number_lineages = 0` we have the lineage-null model (i.e. the environment model, with no cells), when `$number_lineages = 1` we have the single-lineage model, and so on.

In the example we gave earlier, lineages can easily be added and removed to $\mathcal{L}$ because each lineage's equations and parameters are indexed via the ℓ subscript, which acts as `$lineage`, while the size of the $\mathcal{L}$ set acts as `$number_lineages`.

To summarize, the full model definition is

- `$ADModel`
  - `variables`
    - `environment_variables` ($\boldsymbol{\nu}$)
      - $\{\nu_1, \nu_2, \dots\}$
    - `lineage_variables` ($\mathcal{X}$)
      - `lineage_1` ($\boldsymbol{x}_1$)
        - $\{x_{1,1}, x_{1,2}, \dots\}$
      - `lineage_2` ($\boldsymbol{x}_2$)
        - $\{x_{2,1}, x_{2,2}, \dots\}$
      - ...
  - `parameters`
    - `model_parameters` ($\boldsymbol{m}$)
      - $\{m_1, m_2, \dots\}$
    - `adaptation_parameters` ($\mathcal{A}$)
      - `lineage_1` ($\boldsymbol{\alpha}_1$)
        - `adaptation_vector_1` ($\boldsymbol{a_{1,1}}$)
          - $\{a_{1,1,1}, a_{1,1,2}, \dots\}$ ..... on simplex
        - `adaptation_vector_2` ($\boldsymbol{a_{1,2}}$)
          - $\{a_{1,2,1}, a_{1,2,2}, \dots\}$ ..... on simplex
        - ...
      - `lineage_2` ($\boldsymbol{\alpha}_2$)
        - `adaptation_vector_1` ($\boldsymbol{a_{2,1}}$)
          - $\{a_{2,1,1}, a_{2,1,2}, \dots\}$ ..... on simplex
        - `adaptation_vector_2` ($\boldsymbol{a_{2,2}}$)
          - $\{a_{2,2,1}, a_{2,2,2}, \dots\}$ ..... on simplex
        - ...
      - ...
  - `equations`
    - $\dot{x}_{\texttt{\$lineage},1}$
    - $\dot{x}_{\texttt{\$lineage},2}$
    - ..... lineage equations
    - $\dot{\nu}_1$
    - $\dot{\nu}_2$
    - ..... environment equations

This information is sufficient to declare an `$ADModel` object, which is the fundamental unit of the AD software. The software parses the generalized model specification and generates dynamical models from the set

$$\mathcal{M} = \{M_0, M_1\{\ell_1\}, M_2\{\ell_1, \ell_2\}) \dots\}$$

where $M_\emptyset$ is the model with zero lineages (the lineage-null model), $M_1\{\ell_1\}$ the

model with a single lineage, and so on, as required according to the generalized model definition in `$ADModel`. As we simulate the adaptation process it becomes necessary to add new (nascent) lineages and remove lineages that become extinct; the generalized model specification, $\mathcal{M}$, contains the information necessary to modify the current dynamical model in order to accommodate the simulation process.

The software's methods have been loosely divided into a basic core of utility methods, which handle the fundamental simulation components (such as simulating the invasion of a nascent mutant lineage), and an extended set of methods, which synthesize the former to comprehensively model the process of adaptation (such as identifying and simulating all possible invasion models).

3.3 Core numerical methods

The standard methods are presented as software functions and are available for direct use. They require the explicit definition of adaptation parameters and any necessary model parameters. From a work-flow perspective, we use the core methods when designing a new generalized model and when we wish to verify the output from the extended methods; otherwise, most of the software's operation is handled automatically by the extended set.

3.3.1 Model simulations

The most common subsequent operation is to simulate the model using specified or random initial values. Standard numerical methods [124] for solving differential equations are used (implemented in the Wolfram Language [123]) to calculate the solution in a given time interval or until steady state is reached. Steady state detection is implemented by injecting events (inequalities) into the equations which are inspected by the numerical methods. When an event occurs an action is taken - in this case, integration stops. The stationarity criterion may be changed by the user but acceptable results have been obtained with

$$||\dot{\mathcal{X}}(t) + \dot{\boldsymbol{\nu}}(t)|| < \delta$$

which stops integration when the norm of the derivatives becomes smaller than δ. Successful simulation will create an `ADSimulation` object that is linked to the `ADModel` object and contains the system trajectory (or just the steady state, if a large number of simulation outcomes must be kept in memory).

A model with oscillating behavior is unlikely to trigger the stationarity criterion, since the sum of derivatives will not be small (unless all variables oscillate in phase), and will simply run to the maximum number of iterations in the numerical integration procedure. Such an event will be reported to the user, who can then decide how to proceed. While it may be possible that increasing the maximum number of iterations will allow the solution to converge, the default value in the software is set high, and we have not encountered the above scenario in practice.

In general, the software expects numerical solutions that reach steady state and a

limitation of our approach is that it is not equipped to handle models with oscillations and limit cycles. In our applications of the software to models of adaptation dynamics, we have only worked with models where all numerical simulations have converged to a steady state even after extensive parameter scans and a very large number of simulations. We would advise users of the software to explore possible behaviors of their model before committing to large-scale simulations.

3.3.2 Model inspection and resolution

A pair of `ADSimulation` and `ADModel` objects can then be inspected to determine the fate of the microbial lineages in the model; for example, to determine the fate of a nascent lineage with novel adaptation parameters. A set of (resolution) rules must be defined for a consistent resolution of simulation outcomes: it is usually sufficient to define a steady state fractional abundance threshold, ϵ, that lineages must exceed in order to be considered extant.

A new model can then be generated by retaining only the extant lineages. For the model $M_n\{\ell_1, \ell_2, \ldots, \ell_n\}$ with n initially extant lineages $\{\ell_1, \ell_2, \ldots, \ell_n\} \in \mathcal{L}$ we obtain the model transition

$$M_n\{\ell_1, \ell_2, \ldots, \ell_n\} \rightarrow M_{n'}\{\ell'_1, \ell'_2, \ldots, \ell_{n'}\} \tag{3.3}$$

where the right-hand side consists of the model $M_{n'}\{\ell'_1, \ell'_2, \ldots, \ell'_{n'}\}$ with n' extant lineages $\{\ell'_1, \ell'_2, \ldots, \ell'_{n'}\} \in \mathcal{L}'$, which satisfy the survival criterion, $||\boldsymbol{x}^*_i|| > \epsilon$, for all $\ell'_i \in \mathcal{L}'$, at steady state (denoted hereafter by the asterisk superscript).

3.3.3 Propagation and augmentation of dynamical models

Acquiring a model transition (Equation 3.3) is usually enough for validating an analytic result that predicts, for example, the outcome of an invasion event, or for verifying results obtained through the more high-throughput, extended, methods. However, the outcome of a single invasion event is usually a small component in describing the adaptation process; more frequently, we want to investigate the outcome of a succession of adaptation events.

The software allows the user to generate a new model from the `ADModel` object, its

`ADSimulation`, and the inspected model transition. The new model can contain (usually) one or more nascent lineages, with specified adaptation parameters, or randomly sampled parameters from an adaptation kernel (Section 3.4.1). In the language of classical adaptive dynamics theory, the transition's RHS model is "augmented" to include the nascent lineage

$$M_n\{\ell_1, \ell_2, \ldots, \ell_n\} \xrightarrow{inspection} M_{n'}\{\ell'_1, \ell'_2, \ldots, \ell_{n'}\} \xrightarrow{augmentation} M_{n'+1}\{\ell'_1, \ell'_2, \ldots, \ell_{n'}, \ell_{n'+1}\}$$

By convention, the right-most lineages of an augmented model are nascent lineages. Extinct lineages are removed from the list and surviving lineages are 'shifted left' in the lineages list to preserve the order of appearance. Further consistency checks are implemented by the software to ensure lineages are tracked correctly. The steady state values of the extant lineages are propagated to the initial values of the derived model

$$\mathcal{X} \ni \boldsymbol{x}_i(t^*) \to \boldsymbol{x}_{i'}(0) \in \mathcal{X}'$$

along with the associated adaptation parameters

$$\mathcal{A} \ni \boldsymbol{\alpha}_i \to \boldsymbol{\alpha}_{i'} \in \mathcal{A}'$$

The nascent lineage is initially scarce and it is usually sufficient to define a relative fractional abundance at which it is introduced into the model

$$||\boldsymbol{x}_{n'}(0)|| \sim \phi \cdot \min_{\mathcal{X}'}\{||\boldsymbol{x}'_i||\}$$

$\phi = 10^{-3}$ is the default value for the initial fractional abundance in the software and should be assumed in this and the following chapters unless otherwise stated.

3.3.4 Summary

Figure 3.3.1 summarizes the software's core objects and methods in a typical work-flow from model specification, model simulation and resolution, to model augmentation and propagation. First, a model is specified and cast in the form

required by the software – as we have done in Section 3.2 for the structured model with a resource allocation trade-off between biomass synthesis and nutrient uptake. This allows us to define the generalized model object, `$ADModel`, which can be used to generate models with any number of lineages. We can then introduce the first single lineage, ℓ_x; to do this, we specify its adaptation parameter – r_x in our example as well as a small initial concentration. The single-lineage model, $M_1(\ell_x)$, is created and simulated. The steady state is inspected and, supposing the lineage survives, a new model is returned, $M_1(\ell_x)$. We then simulate invasion of a second lineage, ℓ_y, after defining its adaptation parameter – i.e. r_y – and initial concentration. The $M_1(\ell_x)$ model is augmented to a two-lineage model, $M_2(\ell_x, \ell_y)$. After simulating the model, we can inspect the solution to obtain a model with all extant lineages. For example, if lineage ℓ_y did not survive, the software will return $M_1(\ell_x)$. This can then be again augmented with a new lineage, ℓ_z, to simulate a new invasion. This process of iteratively generating/augmenting, simulating, and inspecting models is the fundamental unit of the numerical methods that we will later expand to fully explore the adaptation dynamics of a model.

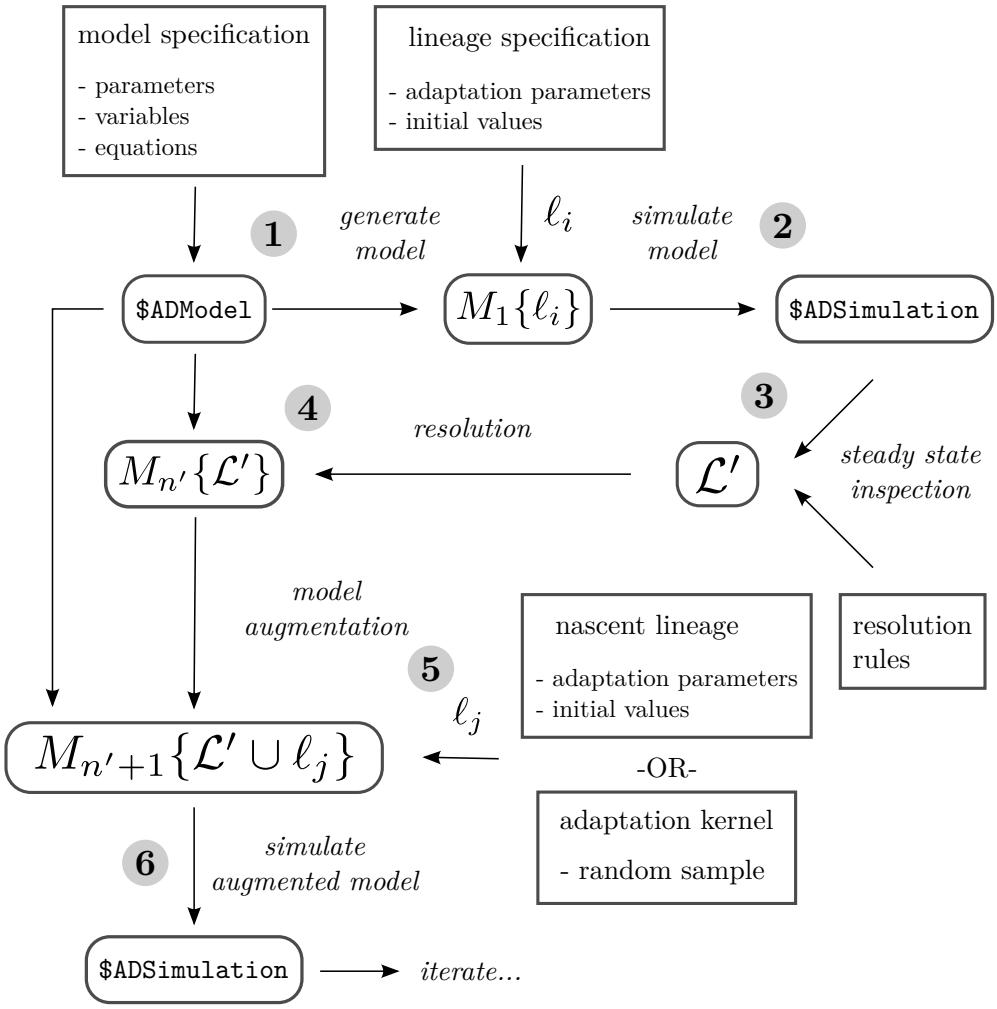


Figure 3.3.1: Summary of a typical work-flow using only the software's core functionality. **1**: The model (rate) equations, variables, and parameters are supplied to the software. This creates the generalized adaptation model, `$ADModel`. A starting lineage, ℓ_i is declared, by specifying its adaptation (i.e. phenotypic) parameters and its initial abundance. The model with a single nascent lineage, M_1, is derived from `$ADModel` using this lineage. **2**: The model is simulated to steady state. **3**: Using a set of provided resolution rules (including extinction thresholds) the `$ADSimulation` is inspected and the set of extant lineages, $\mathcal{L}'$ is obtained. In this simple case, $\mathcal{L}'$ is either empty or just contains ℓ_i, since no other lineages have been introduced yet. **4**: Based on the number of extant lineages, n', a new model is derived from the generalized definition which contains the extant lineages. **5**: The resolved model is augmented with a new, nascent, mutant lineage, ℓ_j, which is either specified explicitly or randomly sampled from an adaptation kernel centered on the resident lineages, $\mathcal{L}'$. **6**: The augmented model is simulated to steady state. This work-flow can then be iterated to explore the adaptation process.

3.4 Extended numerical methods

The software's extended methods automate and build on the core methods described so far to construct a more comprehensive picture of adaptation dynamics. Manual intervention and supervision has been minimized and manual input has been replaced by rules and statistical processes. The main outputs of automatic mode are:

1. Adaptation trajectories and trajectory ensembles
2. Adaptation maps
3. Adaptation graphs

3.4.1 Adaptation trajectories and trajectory ensembles

An adaptation trajectory is a discrete-time representation of a sequence of models, derived from the generalized `ADModel` object, as nascent lineages emerge and compete with extant resident lineages. Trajectories are stored as model transitions

$$T = \{M_{n_0}\{\mathcal{L}_0\} \rightarrow M'_{n_0}\{\mathcal{L}_{0'}\}, M_{n_1}\{\mathcal{L}_1\} \rightarrow M'_{n_1}\{\mathcal{L}'_1\}, \dots\}$$

where n_i denotes the number of extant lineages ($\mathcal{L}_i$) at the beginning of the i-th time step and n'_i the number of extant lineages ($\mathcal{L}'_i$) at the end of the i-th time step, following simulation, steady state inspection, and removal of extinct lineages.

One nascent lineage is generated at the beginning of every time step, whose parameters are sampled from an adaptation kernel centered on the extant resident lineages from the previous time step. In other words, $\mathcal{L}_{k+1} = \mathcal{L}'_k \cup \ell_m$, where ℓ_m is a random, nascent, mutant lineage. Random sampling from the Dirichlet distribution,

$$f(x_1, \dots, x_d) = \frac{1}{B(\boldsymbol{\alpha})} \prod_{i=1}^{K} x_i^{\alpha_i - 1}$$

with $B(\boldsymbol{\alpha})$ the multivariate Beta function and $\alpha = \{1, 1, \dots\}, \alpha \in \mathbb{R}^d$, followed by normalization, the samples can be used to generate points that are randomly distributed on the d-dimensional unit simplex.

The kernel has one parameter, `$max_adaptation_norm`, which is used to limit the maximum (Euclidean) distance between the progenitor resident lineage and the nascent lineage in the adaptation parameter space. This is enforced via naive rejection sampling: keep sampling using $Dir[\alpha]$ until we generate an acceptable point on the unit simplex[1]. When there is more than one extant resident a progenitor is first select by random choice weighted by the fractional abundance of each resident lineage. Finally, it is usually assumed that only one adaptation vector mutates and that only one nascent lineage bearing a mutation appears at every step of the adaptation process.

Trajectories are initialized with the linage-null model, $M_0\{\emptyset\}$, which is then simulated to allow the environment variables to reach steady state without interference from cell populations. The $M_1\{\ell_1\}$ model is then derived from the lineage-null steady state with a nascent ancestral lineage generated by sampling the adaptive parameter spaces without restriction. The $M_1\{\ell_1\}$ model is then simulated to steady state, inspected, and a new nascent lineage is introduced. Simulation stops either when a specified maximum number of time steps has been reached or when a stop criterion is met. By default, the trajectory stops if no lineage's adaptation vector changes by more than σ from its nearest neighbor for a number of maximum time steps, `$max_stationary_steps`:

$$\max_i\{\min_j\{||\alpha_i(t) - \alpha_j(t-1)||\}\} < \sigma \text{ for all pairs } \alpha_i \in \mathcal{A}'(t), \alpha_j \in \mathcal{A}'(t-1)$$

for `$max_stationary_steps`

It is implied that if the number of co-resident lineages changes between t and $t-1$ that the stationarity counter is not incremented, even if the vector norms are smaller than σ.

On successful exit, the trajectory will be saved as a `ADTrajectory` object. The trajectory can be visualized to show the time evolution of specified adaptation parameters, various trajectory properties, like the number of co-residents, and

[1] This is suboptimal, and can cause delays during trajectory sampling when `$max_adaptation_norm` is small, but restricted sampling on the unit simplex is not trivial.

other features of interest, such as co-residency level transitions (bifurcation and convergence points). The complete procedure for generating a single adaptation trajectory is shown in Algorithm 1.

While adaptation dynamics trajectories have much in common with ordinary dynamical trajectories they also possess features that make them more subtle to analyze. Perhaps most importantly, unlike smooth dynamical processes, the trajectories are non-local because system state transitions are not always continuous (except, approximately, for very small maximum mutation norms; in addition, the dimensionality of the system state (the resident phenotypes) can change as a result of additional lineages establishing in the environment and lineages becoming extinct.

A trajectory ensemble can be constructed by sampling multiple independent trajectories, and software methods are provided for parallel computation of ensembles of independent trajectories. It is possible to specify unique simulation parameters for each trajectory, if needed. This can be useful if trajectories are required for the same `ADModel`, but with different maximum adaptation (mutation) norms, in order to investigate adaptation dynamics at a 'local' and at a'global' scale. A group of adaptation trajectories is stored in an `ADEnsemble` object as an array of `ADTrajectory` objects together with associated simulation parameters.

Trajectory ensembles can be used to construct adaptation graphs, which provide a more comprehensive view of adaptation dynamics and are discussed in Section 3.4.3.

3.4.2 Adaptation maps and trees

A more systematic approach to exploring adaptation space than sampling random trajectories is to simulate all possible models from $\mathcal{M} = \{M_0, M_1\{\ell_1\}, M_2\{\ell_1, \ell_2\}) \dots\}$, for all possible adaptation parameter values and co-residency levels. Since the adaptation space is continuous this is, of course, impossible; a compromise, then, is to use a discretized adaptation space and work only with the models permitted on that space.

Discretization is performed by taking u uniformly spaced points on the unit line. For a $d-$dimensional adaptation vector, the associated discretized $d-$dimensional unit simplex will have

Algorithm 1 Adaptation trajectories

Require: `adaptation kernel (parameters)`, `stop criteria`

Initialize `$trajectory` as empty array (dynamic).
Initialize `$currentState` $= M_0\{\emptyset\}$
Simulate and inspect $M_0\{\emptyset\} \to M_0\{\emptyset\}$ ▷ RHS at steady state
`$trajectory` $\leftarrow (M_0\{\emptyset\} \to M_0\{\emptyset\})$
while `stop criteria` $== False$ **do**
    Set `$previousTransition` $= (M_{n_i}\{\mathcal{L}_n\} \to M_{n'_i}\{\mathcal{L}_{n'}\})$
    Get nascent lineage : $\ell_j =$ `GenerateNascentLineage`$(M_{n'_i}\{\mathcal{L}'_i\})$
    Derive new model, simulate, and inspect

$$M_{n'_i}\{\mathcal{L}'_i\} \xrightarrow{\alpha_j} M_{n_i+1}\{\mathcal{L}'_i \cup \ell_j\} = M_{n_{i+1}}\{\mathcal{L}_{i+1}\} \to M_{n'_{i+1}}\{\mathcal{L}'_{i+1}\}$$

    Append `$trajectory` $\leftarrow \left(M_{n'_i}\{\mathcal{L}'_i\} \to M_{n'_{i+1}}\{\mathcal{L}'_{i+1}\}\right)$
end while
return `$trajectory`

function GENERATENASCENTLINEAGE($M_n\{\mathcal{L}\}$)
    Choose random progenitor lineage, ℓ_i, from the n lineages with weights $\{||\boldsymbol{x}_1||, ||\boldsymbol{x}_2||, \ldots, ||\boldsymbol{x}_n||\}$, and adaptation tuple α_i.
    Choose random adaptation vector to modify, $\boldsymbol{a}_{i,k}$.
    Sample and normalize $\boldsymbol{a}_{j,k} = \dfrac{\tilde{\boldsymbol{a}}_{j,k}}{||\boldsymbol{a}_{j,k}||} \sim Dir[\alpha]$ ▷ $\alpha = \mathbf{1}$
    while $||\boldsymbol{a}_{i,k} - \boldsymbol{a}_{j,k}|| >$ `$maxNorm` **do**
        Sample and normalize $\boldsymbol{a}_{j,k} = \dfrac{\tilde{\boldsymbol{a}}_{j,k}}{||\boldsymbol{a}_{j,k}||} \sim Dir[\alpha]$ ▷ $\alpha = \mathbf{1}$
    end while
    return Lineage ℓ_j with modified adaptation parameter, $\boldsymbol{a}_{j,k}$
end function

$$q = \binom{u + (d-1)}{d-1}$$

points, which are uniformly spaced apart. Assuming there is only one adaptation vector, and therefore only one simplex, each point on the discretized simplex represents the possible adaptation parameters for a single lineage that can exist in `ADModel`.

The objective then is to simulate all models from $\mathcal{M}$ on the discretized simplex, $\tilde{S}$. The naive brute-force approach is to, for each $M_n\{\mathcal{L}\}$ permitted on $\tilde{S}$, to derive the $M_{n+1}\{\mathcal{L} \cup \ell_j\}$ model, for each $\ell_j \in \tilde{S}$, simulate it, and inspect the result to construct a transition. An overview of the procedure is shown in Algorithm 2 and a visual representation is shown in Figure 3.4.1.

Algorithm 2 Adaptation maps (naive)

Require: Maximum co-residency level $R_{max} \geq 0$

1: Initialize empty `$ADMap` array (dynamic).
2: **for** co-residency level $R = 0$ to R_{max} **do**
3: **for** $\mathcal{A}_i$ in all length-R tuple combinations from the $\alpha_i \in \tilde{S}$ **do**
4: **for** $\alpha_j \in \tilde{S}$ & $\alpha_j \notin \mathcal{A}_i$ **do**
5: Augment: $M_R\{\mathcal{L}_i\} \xrightarrow{\alpha_j} M_{R+1}\{\mathcal{L}_i \cup \ell_j\}$
6: Simulate and inspect steady state: $M_{R+1}\{\mathcal{L}_i \cup \ell_j\} \xrightarrow{inspect} M_{n'}\{\mathcal{L}'\}$
7: Save transition

$$\texttt{\$ADMap} \leftarrow \left(M_R\{\mathcal{L}_i\} \xrightarrow{\alpha_j} M_{R+1}\{\mathcal{L}_i \cup \ell_j\} \xrightarrow{inspect} M_{n'}\{\mathcal{L}'\}\right)$$

8: **end for**
9: **end for**
10: **end for**
11: **return** `$ADMap`

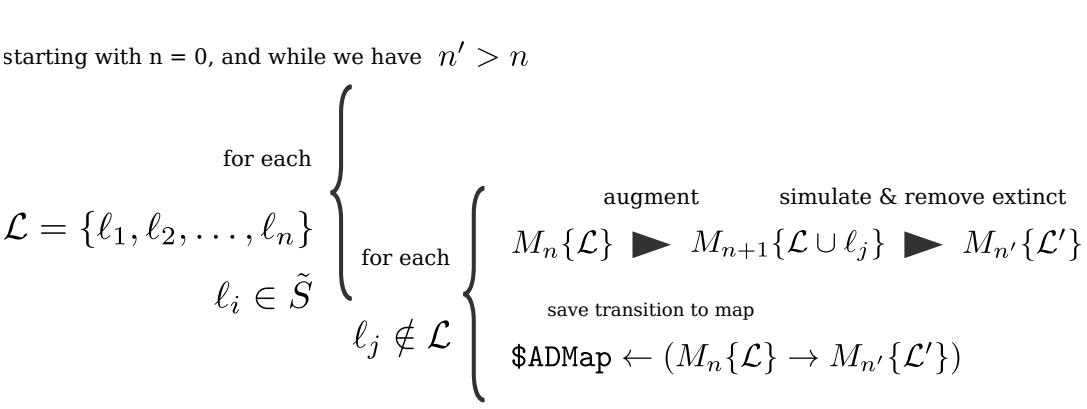


Figure 3.4.1: Overview of the naive mapping algorithm, which simulates all possible invasions for every possible model on the discretized simplex. The objective is to return the `$ADMap` object, which contains all possible model transitions under mutation/invasion events. We start with the model with no lineages, M_0, which contains only environmental (e.g. nutrient) variables. We simulate the model to steady state and, for each possible phenotypic lineage on the discretized simplex, $\ell_j \in \tilde{S}$, create an augmented model, $M_1\{\ell_j\}$. We simulate the augmented model and inspect the steady state to remove any extinct lineages to keep $M_{n'}\{\mathcal{L}'\}$. The transition implied by this mutation/invasion event is saved in the `$ADMap` object. When we have finished iterating through all the ℓ_j on the simplex, we repeat the process but with $M_1\{\ell_i\}$ as the starting model that we will later augment; however, since we have many $\ell_i \in \tilde{S}$, we need to do this for every possible initial model. The mapping algorithm stops when no simulation generates a model with a co-residency level larger than the current one – i.e. $n' \leq n$. This guarantees that we have explored all possible mutation/invasion models on the discretized simplex.

A more efficient approach, using principles of dynamic programming, is to construct a tree of model transitions, with $M_0\{\emptyset\}$ as the root node. Nodes (models) are added to the tree in a breadth-first manner, and simulations are only carried out when needed - i.e. if the inspected transition does not already exist in the tree. An overview of the procedure is shown in Algorithm 3 and an illustration for a simple case is shown in Figure 3.4.2.

The output of the dynamic programming algorithm is an `$ADTree` object, which is a more efficient form of the `$ADMap` object, that contains all possible transitions between all possible dynamically stable models.

Algorithm 3 Adaptation maps (dynamic)

Initialize an empty `$ADTree` tree data type ▷ Recursive
Initialize (FIFO) queue `$active_nodes` with $M_0\{\emptyset\}$ as first member
while $|$`$active_nodes`$| > 0$ **do**
    Dequeue front queue member, $M_n\{\mathcal{L}\}$ ▷ Parent node
    for $\alpha_j \in \tilde{S}$ & $\alpha_j \notin \mathcal{L}$ **do** ▷ Child nodes
        if $\left(M_n\{\mathcal{L}\} \xrightarrow{\alpha_j} M_{n+1}\{\mathcal{L} \cup \ell_j\} \to M_{n'}\{\mathcal{L}'\}\right) \in$ `$ADTree` **then**
            Copy entry to tree

$$\texttt{\$ADTree} \leftarrow \left(M_n\{\mathcal{L}\} \xrightarrow{\alpha_j} M_{n+1}\{\mathcal{L} \cup \ell_j\} \to M_{n'}\{\mathcal{L}'\}\right)$$

        else
            Derive $M_n\{\mathcal{L}\} \xrightarrow{\alpha_j} M_{n+1}\{\mathcal{L} \cup \ell_j\}$
            Simulate derived model and inspect $M_{n+1}\{\mathcal{L} \cup \ell_j\} \to M_{n'}\{\mathcal{L}'\}$
            Save entry to tree

$$\texttt{\$ADTree} \leftarrow \left(M_n\{\mathcal{L}\} \xrightarrow{\alpha_j} M_{n+1}\{\mathcal{L} \cup \ell_j\} \to M_{n'}\{\mathcal{L}'\}\right)$$

            Enqueue $M_{n'}\{\mathcal{L}'\}$ to the back of `$active_nodes`
        end if
    end for
end while
return `$ADTree`

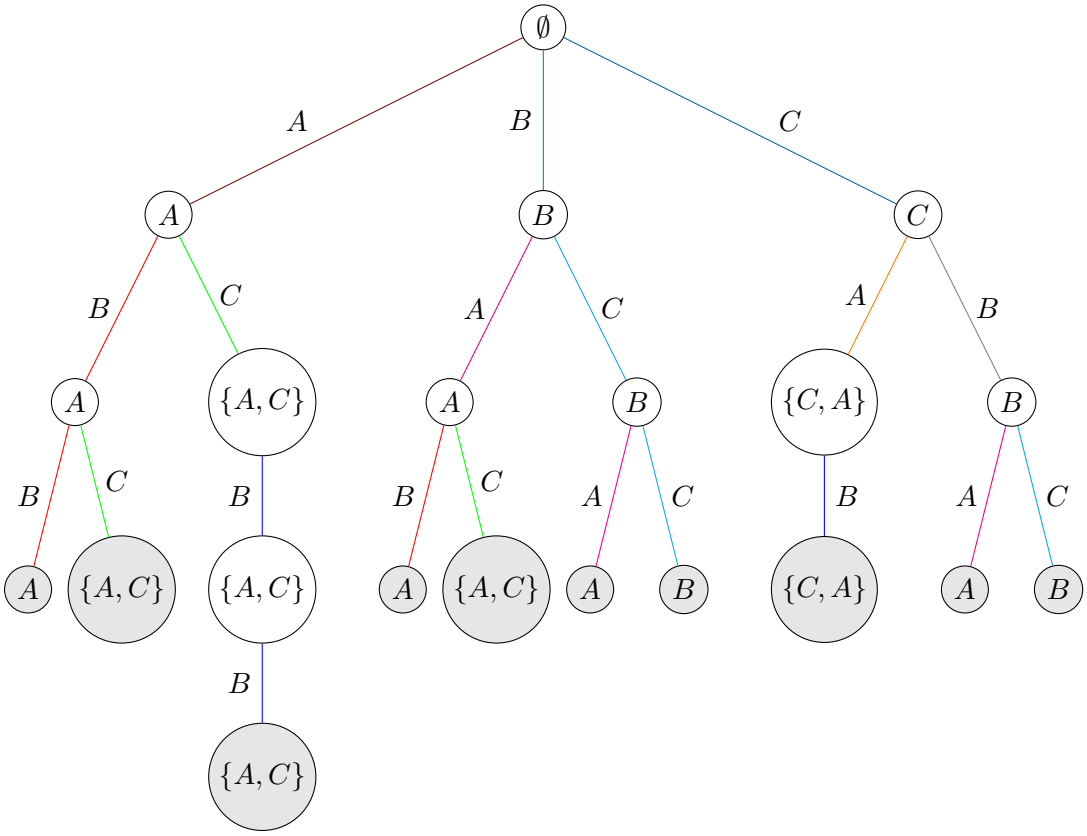


Figure 3.4.2: The dynamic mapping algorithm applied to a simple adaptation process example. Nodes in the tree represent models and edges represent transitions from an ancestral model (parent node) to a new state (child node) after augmented the resolved ancestral model with a nascent lineage. In this example, the discretized simplex, $\tilde{S}$, has only three possible adaptation parameter values: A, B and C. Nodes are labeled with the adaptation parameters of their extant lineages and edges are labeled with the nascent lineage that augments the parent node. Light nodes are models that have passed through the simulation queue while dark nodes are models that have not been through this queue. Model transitions (edges) have been given unique colors to illustrate why and when the algorithm avoids redundant simulations. The algorithm is initialized with the lineage-null model ($\emptyset$) in the queue. Models are dequeued from the 'front' of the queue (i.e. in a breadth-first tree traversal) and their child nodes are populated by augmenting the (parent) model ($\mathcal{A}$) with every possible adaptation parameter ($\alpha_i \in \tilde{S}, \alpha_i \notin \mathcal{A}$). Models are simulated if they do not already exist in the tree: if they do, the steady state outcome is copied from the existing result, and the child node is placed appropriately in the tree; if the model has never been simulated, then it will be simulated to steady state and resolved, and its child nodes (all possible augmented models) will be placed on the tree and enqueued to the back of the queue. The algorithm stops when there are no more models in the queue.

3.4.3 Adaptation graphs

A natural way to visualize adaptation dynamics for `$ADModel` objects is to construct graphs where the vertices are dynamically-stable models ($M_n\{\mathcal{L}\}$). Two vertices are connected if a nascent lineage can change the state (lineage/phenotype composition) of the source vertex to that of the target vertex. Moreover, graphs can be analyzed as discrete Markov processes to determine process properties, such as the long-term (stationary) vertex distributions, which are the end-points of adaptation trajectories.

There are two software methods for constructing adaptation graphs: from adaptation maps and from trajectory ensembles.

3.4.3.1 Graph construction from adaptation maps

Constructing the graph of an adaptation map is straightforward when the `$ADTree` data structure is used:

1. Each node in the tree with lineage composition $\{\ell_1, \ell_2, \ldots, \ell_n\}$ becomes a unique vertex in the graph. Lineage composition is treated as an unordered set for determining uniqueness, therefore the order in which lineages appear is not important. The lineage-null model is omitted to produce an unrooted graph.

2. Each edge between a parent and child node in the tree becomes an edge between the corresponding vertices in the graph.

3. Duplicate edges and self-edges may be removed for styling purposes and when the simple graph, instead of the multi graph, is required.

The example shown in Figure 3.4.2 has the (simple) graph shown in Figure 3.4.3.

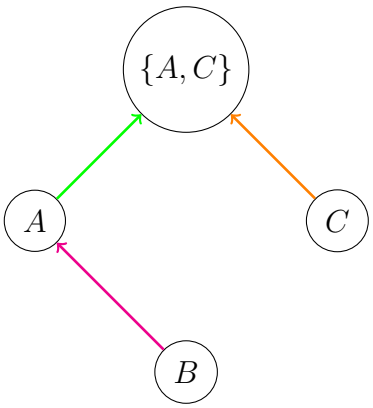


Figure 3.4.3: The (simple) graph of the adaptation map example shown in Figure 3.4.2. Edge color has been preserved between the two figures. Self-edges and repeated edges have been removed.

Algorithm 4 Graph construction from adaptation maps

Require: `$ADTree` or `$ADMap` object

1: Initialize `$nodes` and `$edges`
2: **for** each entry $\left(M_n\{\mathcal{L}\} \xrightarrow{\alpha_j} M_{n+1}\{\mathcal{L} \cup \ell_j\} \to M_{n'}\{\mathcal{L}'\}\right)$ **do**
3: **if** $M_n\{\mathcal{L}\} \notin$ `$nodes` **then** ▷ Treating $\mathcal{L}$ as unordered set
4: `$nodes` $\leftarrow M_n\{\mathcal{L}\}$
5: **end if**
6: **if** $M_{n'}\{\mathcal{L}'\} \notin$ `$nodes` **then** ▷ Treating $\mathcal{L}'$ as unordered set
7: `$nodes` $\leftarrow M_{n'}\{\mathcal{L}'\}$
8: **end if**
9: `$edges` $\leftarrow (M_n\{\mathcal{L}\} \to M_{n'}\{\mathcal{L}'\})$ ▷ No duplicates for simple graph
10: **end for**
11: **return** Graph object, G(`$nodes`, `$edges`)

Note that graphs constructed in this way are unrestricted: all parent-child connections are included in the graph regardless of the adaptation parameters of the nascent lineage. In the discrete Markov process analysis of adaptation graphs, which is described in Section 3.4.4, it is possible to restrict parent-child connections by specifying an adaptation kernel or maximum mutation norm to truncate the graph; for example, restrict graph edges to transitions mediated by nascent lineages with adaptation parameters similar to those of their resident progenitors.

Constructing adaptation graphs from `$ADMap` objects has two main drawbacks: one related to algorithm complexity, and one related to interpretation.

First, complexity grows in proportion to the number of simplex elements, $q = \binom{u+(d-1)}{d-1}$, raised to the power of the maximum possible co-residency level, R_{max}, in `$ADModel`.

Complexity is of order

$$\mathcal{O}\left(\binom{u+(d-1)}{d-1}^{R_{max}+1}\right)$$

To be more precise, the naive mapping algorithm will execute exactly

$$\sum_{R=0}^{R_{max}}\left(\binom{q}{R} * (q-R)\right) \tag{3.4}$$

model simulations, whereas this is only the worst-case complexity for the dynamic algorithm (see Figure 3.4.4 for indicative numbers). We can derive expression 3.4 inductively:

- At co-residency level 0 there are q models to derive, one for each of the q nascent lineage with distinct adaptation parameters.
- At co-residency level 1 there are at most $q = \binom{q}{1}$ unique models (if all q lineages from the previous level have fixed in the environment). For each of these models there are $q-1$ possible (non-duplicate) nascent lineages with which to augment and hence there are $q(q-1)$ models to simulate.
- At co-residency level 2 there are at most $\binom{q}{2}$ unique models, after accounting for lineage re-arrangements, and each model can be augmented $q-2$ times with non-duplicate nascent lineages.
- At the highest co-residency level possible, R_{max}, there are at most $\binom{q}{R_{max}}$ established models; therefore, for each model, there are $q - R_{max}$ possible augmentations.

The second drawback in producing graphs of adaptation dynamics from maps is the difficulty in interpreting the results. Because it is not possible to simulate every possible model with continuous adaptation parameters it was necessary to first discretize the unit simplex(es) and then construct the map of all possible models. However, the discretization process may introduce 'gaps' to adaptation

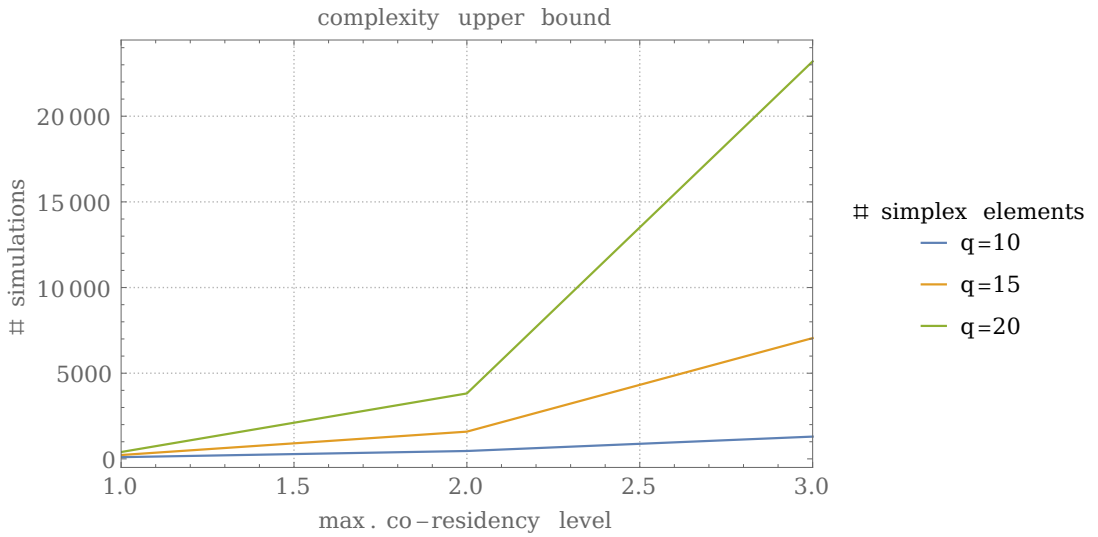


Figure 3.4.4: Mapping algorithm complexity, in terms of numbers of simulations required to produce a complete map. The plot shows the worst-case complexity (upper bound) for the dynamic algorithm for the number of simplex elements and the maximum co-residency level permitted on the model. The naive mapping algorithm will always simulate the indicate number of models while the dynamic algorithm will typically require approximately half of this number for a complete map.

dynamics that would not otherwise be present in the continuous formulation. See Figure 3.4.5 for an illustration of the problem. The implication is that the discretization scheme chosen for a particular `$ADModel` will affect the adaptation graph, and, consequently, the discrete process analysis - not least, the stationary vertex distribution.

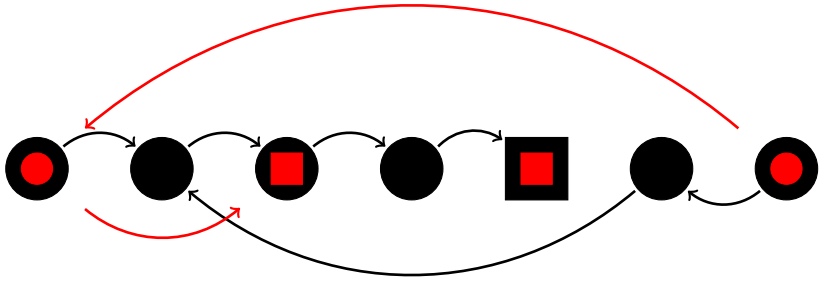


Figure 3.4.5: Discretization barriers. In this illustration, the adaptation space has been discretized either to 7 points (black nodes) or 4 points (red nodes). The edges show node transitions. For the red nodes (4 points) only the red edges will be discovered by the mapping algorithm, while for the black nodes (7 points) both red and black edges will be discovered. Square nodes represent stationary states (no edge out of the node) in the map while circles are non-stationary (at least one edge out of the node). There are two stationary nodes for 4-point discretization (red #2 and red #3, numbering from the left) but only one stationary node (black #5, numbering from the left) for 7-point discretization because the intermediate node (black #4) required to access the black stationary node is not present in the 4-point discretization.

3.4.3.2 Construction from trajectory ensembles

Constructing graphs from trajectory ensembles is a more nuanced task compared to constructing graphs from adaptation maps; however, it is possible to overcome the discretization 'gap' problem by exploring the continuous parameter space first, and discretizing later; moreover, it may be possible to do so using a procedure with better complexity than the mapping algorithms, at least in some cases.

Starting with an `$ADEnsemble` object (Section 3.4.1) the objective is to analyze the individual trajectories, and the ensemble as a whole, and construct the nodes and edges of the graph that best represent the models, and transitions between models, of the ensemble. By necessity, this implies grouping (binning) models, and possibly transitions, based on certain rules and conditions. The choice of grouping strategy will determine which ensemble properties are important and how they are represented in the final graphs.

A general procedure has been implemented in the software, with modularity being one of the emphasized design goals. Constructing the adaptation graph of an ensemble of adaptation trajectories is an optimization problem where each

iteration consists of three parts:

1. Choose a discretization (binning) scheme (simplex mesh) and discretize the trajectories in the ensemble.
2. Construct a candidate graph with nodes representing discretized models (models with binned adaptation parameters) and edges representing transitions between the discretized models.
3. Score the graph using specified graph measures. Accept or reject the discretization scheme based on the score.

Each of the three iterated objectives can be thought of as a separate procedure that can be modified independently of the others to assemble a graph reconstruction algorithm. Finding the 'optimal' graph that represents the salient features of an adaptation model is not an unambiguous task, however, and is contingent on the definition of optimality.

One implementation in particular, which has produced acceptable results in practice and is the software default, is presented in Algorithm 5 and explained below. The procedure seeks to produce the graph that minimizes graph reciprocity while maximizing the fraction of self-edges:

- Reciprocity (Figure 3.4.6) is defined as the fraction of edges that are reciprocated in the simple graph. Two edges are reciprocated if for the (unidirectional) edge between two nodes $\nu_i \to \nu_j$ there exists the edge $\nu_j \to \nu_i$.
- The fraction of self-edges (Figure 3.4.7) is the ratio of the number of edges $\nu_i \to \nu_i$ to all edges in the multi graph.

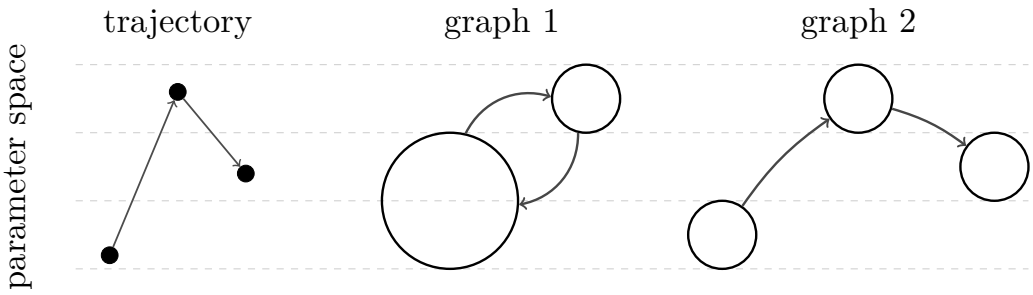


Figure 3.4.6: Reciprocity. In this example, the trajectory from which the graph is to be constructed is shown on the left. In the first graph, a large bin was chosen for the lower part of the parameter space, which encompassed the first and third point. A smaller bin was chosen for the top part of the parameter space, which only included the second point. Because binned vertices inherit the edges of their member points, the resulting graph would have reciprocity = 2/2 - that is, two out of two edges are reciprocated. In the second graph, smaller bins were chosen for the bottom, middle, and upper parts of the parameter space. This choice of bins has removed the reciprocated edges: reciprocity = 0/2.

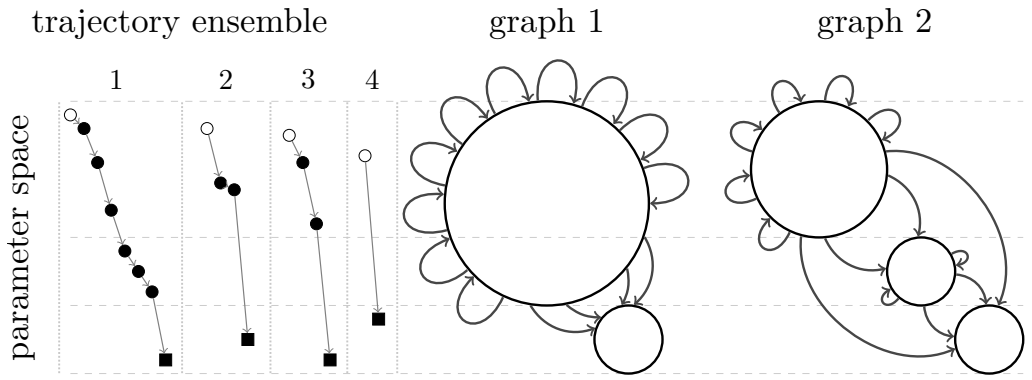


Figure 3.4.7: Self-edges. In this example, there are four trajectories from which to reconstruct the graph. Open circles are starting points, filled circles are intermediate points, and filled squares are final points. In the first graph, a large bin was chosen for the top part of the parameter space and a small one for the bottom part. The ratio of self-edges to all edges is 10/14. In the second graph, the larger bin was divided (unequally) into two smaller bins. The new graph has a self-edge ratio of 8/14.

Minimizing reciprocity is desirable when we know a priori that the adaptation models we are working with must exclude the possibility of a transition from $B \to A$ if the transition $A \to B$ already exists. This is true when the models are deterministic and when the nascent lineage always has small initial values relative to the ancestral residents. It will also be true when the models do not have frequency-dependent growth rates. The case of stochastic models, and nascent lineages invading above scarcity, or models with strong frequency dependence is more subtle and has not been addressed. Because binned vertices (models) inherit the edges of their constituent models, it will often be true, particularly when using coarse bins, that models with reciprocating transitions will be grouped together. Partitioning the binning interval, then, might reduce reciprocity.

Maximizing the fraction of self-edges is desirable because we can minimize the number of unique vertices required to represent the adaptation dynamics for a particular model. A larger number of self-edges means that the choice of binning intervals is optimal because models with similar behaviors (transitions) are grouped together.

The two optimization objectives are at odds, however. Reciprocity is guaranteed to be 0 when each model in the ensemble is the unique member of a unique bin (with boundary exceptions); however, using the finest possible binning intervals, the fraction of self-edges will be very small(in fact, it will be 0 if we exclude boundary cases and end-points), which means that the graph representation is inefficient. At the other extreme, the coarsest possible interval $[0, 1]$ (in all dimensions), has a self-edge fraction $= 1$, since every edge is a self-edge; however, since self-edges are by definition reciprocated edges, the graph reciprocity is maximum (1).

The 'optimal' graph, therefore, lies between the finest and coarsest discretization scheme. The scoring function used in the default algorithm is a simple weighted linear sum

$$S(\mathcal{G}) = c_\sigma \sigma(\mathcal{G}) + c_\epsilon \epsilon(\mathcal{G})$$

where σ and ϵ are the fraction of self-edges and graph reciprocity respectively, and c_σ and c_ϵ their respective weights. The latter pair is a parameter for the reconstruction algorithm with default values $c_\sigma = 1, c_\epsilon = -1$.

The optimization search follows a naive depth-first interval bisection pattern to find candidate graphs that maximize the graph score (Figure 3.4.8). Because multiple discretization schemes may be equally good, particularly for smaller graphs, multiple candidates may be returned.

The optimization search is initialized with a single element on the stack,

$$\mathcal{G}(G, \mathcal{B}, S) \tag{3.5}$$

This element consists of a candidate graph, G, a discretization scheme (bins), $\mathcal{B}$, and a score, S. The initialization element has the coarsest possible discretization scheme: the interval $[0, 1]$. This will be used to bin model adaptation parameters. Note that the binning intervals are assumed to be symmetric for all dimensions, so an adaptation vector in R^d will have the same binning intervals along each of the d dimensions.

While there are items left on the stack, the top graph element is accessed. The trajectory ensemble is discretized according to this element's discretization scheme ($\mathcal{B}$), and a graph (G) is constructed. The graph is analyzed and given a score (S) according to the scoring function. If this graph scores better than its parent, then its child nodes will be populated and pushed to the top of the stack. One child node for every possible interval bisection is added to the tree and connected to the current node. For example, if the current node has intervals $\mathcal{B} = \{[0, 0.5], [0.5, 0.75], [0.75, 1]\}$ then three child nodes are possible:

- $\mathcal{B}_{child} = \{[0, 0.25], [0.25, 0.5], [0.5, 0.75], [0.75, 1]\}$
- $\mathcal{B}_{child} = \{[0, 0.5], [0.5, 0.625], [0.625, 0.75], [0.75, 1]\}$
- $\mathcal{B}_{child} = \{[0, 0.5], [0.5, 0.75], [0.75, 0.8625], [0.8625, 1]\}$

If, on the other hand, the current node has a score that is no better than the score of its parent, then its child nodes will not be populated and search along this branch of the tree will be aborted. Note that this implies that the algorithm cannot escape local maxima. An attempt is made to make up for this possibility at the end of the optimization search (explained later).

Moreover, if the current node has a score that is equal to the score of the best graph objects constructed so far - anywhere in the tree - then it is added to the

list of best candidate graphs. If the current node has a score that is *higher* than the score of the best graph objects, then the list of best graphs is reset and the current node becomes its only member. In practice, allowing some tolerance for adding elements to the list of best graphs, and reseting this list, has yielded better results and this is allowed in the software.

The search ends when no child can improve the score of its parent - i.e. the stack is empty. Depending on the size of the ensemble and on the weights chosen for reciprocity and self-edges in the scoring function, the search might encounter small score increases for each new candidate; for this reason, a `$maxDepth` parameter is allowed which limits the tree depth at which new nodes can be added.

Algorithm 5 Graph construction (ensembles)

Require: `PopulateChildNodes, DiscretizeConstructGraph, ScoreGraph`

Initialize (LIFO) stack of graph objects, `$graphStack`, with the root element $\mathcal{G}_0\{G_0 = \emptyset, \mathcal{B}_0 = \{[0, 1]\}, S_0 = \emptyset\}$
Initialize the empty collection `$bestGraphs`
Initialize $S_{best} = \emptyset$
while |`$graphStack`| > 0 **do**
    Pop top stack item : $\mathcal{G}\{G = \emptyset, \mathcal{B}, S = \emptyset\}$
    Discretize ensemble, construct graph, and update graph object: $\mathcal{G}(G, \mathcal{B}, S = \emptyset) \leftarrow$ `DiscretizeConstructGraph`(`$ADEnsemble`, $\mathcal{B}$)
    Score graph and update graph object $\mathcal{G}(G, \mathcal{B}, S) \leftarrow$ `ScoreGraph`(G, `$reciprocatedWeight, $selfEdgesWeight`)
    if $S > S_{parent}$ and `$nodeDepth` < `$maxDepth` **then**
        Populate child nodes and push to top of stack `$graphStack` $\leftarrow$ `PopulateChildNodes`($\mathcal{G}$)
    end if
    if $S > S_{best}$ **then**
        Reset `$bestGraphs` $= \{\mathcal{G}\}$.
        Set $S_{best} = S$.
    else if $S == S_{best}$ **then** ▷ Or allow ± tolerance
        Append `$bestGraphs` $\leftarrow \mathcal{G}$
    end if
end while
return `$bestGraphs` ▷ Or other finalizing procedure

function PopulateChildNodes($\mathcal{G}_{parent}(G_{parent}, \mathcal{B}_{parent}, S_{parent})$)
    Initialize empty `$childNodes` of graph objects
    for each `$bin` $\in \mathcal{B}_{parent}$ **do**
        $\mathcal{B}_{child} = (\mathcal{B}_{parent} \setminus$ `$bin`$) \cup$ (`$bin[0]`, $\frac{\texttt{\$bin[0]}+\texttt{\$bin[1]}}{2}$) $\cup$ ($\frac{\texttt{\$bin[0]}+\texttt{\$bin[1]}}{2}$, `$bin[1]`) ▷ New bins with bisected `$bin`
        `$childNodes` $\leftarrow \mathcal{G}_{child}(\emptyset, \mathcal{B}_{child}, \emptyset)$
    end for
    return `$childNodes`
end function

function DiscretizeConstructGraph(`$ADEnsemble`, `$bins`)
    Initialize empty collection of nodes `$nodes` ▷ Preferably a hashed set
    Initialize empty list of edges `$edges`
    for each trajectory $T \in$ `$ADEnsemble` **do**
        for $k = 1$ to $k = |T| - 1$ **do**
            Bin adaptation parameters, $\mathcal{A}'_k$ and $\mathcal{A}'_{k+1}$, of the extant models, $M_{n_k}\{\mathcal{L}_k\}$ and $M_{n_{k+1}}\{\mathcal{L}_{k+1}\}$, in transitions t_k and $t_{k+1} \in T$ using the bins $\{\boldsymbol{B}_p\} \in$ `$bins`: $\boldsymbol{B}_p \leftarrow \mathcal{A}'_k$ and $\boldsymbol{B}_q \leftarrow \mathcal{A}'_{k+1}$
            `$nodes` $\leftarrow \boldsymbol{B}_p, \boldsymbol{B}_q$ ▷ If not already present
            `$edges` $\leftarrow (\boldsymbol{B}_p \rightarrow \boldsymbol{B}_q)$ ▷ Allow duplicates
        end for
    end for
    Construct graph G with `$.nodes` and `$.edges`.
    return Graph object $\mathcal{G}(G, \mathcal{B} =$ `$bins`$, S = \emptyset)$
end function

function ScoreGraph(`$Graph`, `$reciprocatedWeight`, `$selfEdgesWeight`)
    Initialize `$reciprocatedEdges` $= 0$
    Initialize `$selfEdges` $= 0$
    for each edge $(\nu_i \rightarrow \nu_j) \in$ `$Graph.$edges` **do**
        if $(\nu_j \rightarrow \nu_i) \in$ `$Graph.$edges` **then**
            `$selfEdges`$+ = 1$
        end if
        if $\nu_i == \nu_j$ **then**
            `$reciprocatedEdges`$+ = 1$
        end if
    end for
    `$fractionReciprocated` $=$ `$reciprocatedEdges`$/|$`$Graph.$edges`$|$
    `$fractionSelfEdges` $=$ `$selfEdges`$/|$`$Graph.$edges`$|$
    return `$score` $=$ `$fractionReciprocated` $\cdot$ `$reciprocatedWeight` $+$ `$fractionSelfEdges` $\cdot$ `$selfEdgesWeight`
end function

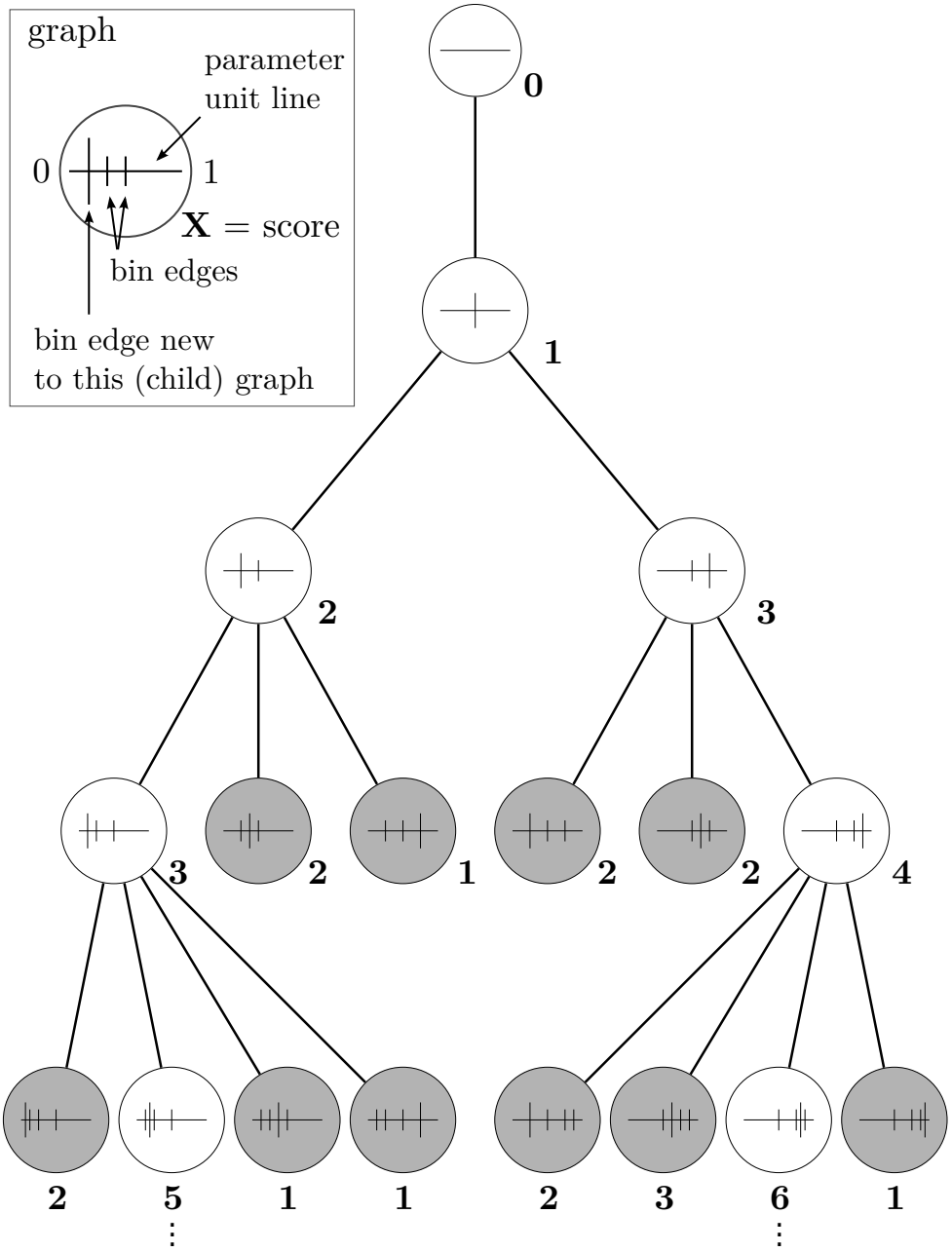


Figure 3.4.8: Algorithm for graph construction from trajectory ensembles applied to an example tree traversal. Nodes are graph objects, $\mathcal{G}(G, \mathcal{B}, S)$ as described in Equation 3.5. The bin edges for each graph object are shown inside the node (see inset for explanation). Light nodes are graph objects that have been through the stack, because their score (bold number next to node) was strictly larger than that of their parent, while gray nodes have not been through the stack. The optimization tree is traversed depth-first and traversal stops when no child node can improve the score of its parent. A list of best-scoring graph objects is retained during traversal and returned at the end. A maximum node depth may be defined to restrict computation time.

In practice, there are usually several 'best' graphs that are returned by the graph construction algorithm. This is an expected consequence of the depth-first bisection, since one branch of the tree will typically refine the discretization scheme in only one part of the parameter space until it finds a local minimum. The default behavior of the software, upon receiving multiple candidate graphs from the reconstruction algorithm, is to produce one final graph by combining the best graphs' discretization schemes.

Algorithm 6 Combine candidate graphs

Require: `$bestGraphs`

1: Initialize `$combinedEdges` = [0.]
2: Set `$currentEdge` = 0.
3: **while** `$currentEdge` $< 1.$ **do**
4: Set `$currentEdge` $= \min_{\texttt{\$bin} \in \texttt{\$bestGraphs}} \{\texttt{\$bin[1]}\}$
5: Append `$combinedEdges` $\leftarrow$ `$currentEdge`
6: **end while**
7: **return** `$combinedEdges`

The graph objects are styled to indicate, amongst other features: the co-residency level of the vertex (number of dynamically stable co-resident lineages), the classification of vertices into stationary or transient (if the stationary distribution has been calculated), and the norm of the mutant phenotype from that of its parent. The graphs are drawn using a variety of standard graph layout algorithms [125] that highlight various aspects of graph structure, such as vertex reachability and clustering.

3.4.4 Stochastic process analysis using Markov chains

The graph of adaptation dynamics is the representation of a finite-state and (homogeneous) finite-time stochastic process. The vertices of the graph are the states of the process and correspond to dynamically-stable lineage compositions, which are characterized by their adaptation parameters. The edges of the graph are transitions and correspond to invasions by nascent lineages with novel adaptation parameters. Let

$$\mathcal{V}_n, \quad n = 0, 1, 2, \ldots$$

be the stochastic process of adaptation with state space $\tilde{\boldsymbol{S}} \in \mathbb{R}^{|\boldsymbol{\alpha}|R_{max}}$. The state space is the discretized unit simplex (or union of simplexes, if there are more than one adaptation vectors), R_{max} is the highest co-residency level, and $|\boldsymbol{\alpha}|$ is the total length of all adaptation vectors.

In these models of adaptation dynamics the novel adaptation parameters are drawn from a kernel centered on the resident lineages' own adaptation parameters. Because the models are also deterministic, the stochastic process also has the Markov property - i.e. the one-step transition probabilities are

$$p_{i,j} = P\{\mathcal{V}_{n+1} = j | \mathcal{V}_n = i\}$$

from state i to state j. Time is interpreted as the number of nascent lineages that have emerged since the ancestral lineage.

The first step in our Markov process analysis is contingent on whether we are dealing with a graph constructed from an `$ADMap` object or an `$ADTree` object (Sections 3.4.2 and 3.4.3.1), or with a graph constructed from an `$ADEnsemble` (Sections 3.4.1 and 3.4.3.2). The reason for this distinction is that the former object is kernel-agnostic, and will require the specification of an adaptation kernel, whereas the latter object originally required a kernel for trajectory sampling.

Maps and trees

To construct the $\mathcal{V}_n$ process for `$ADMap`, first create the binary transition matrix, $\boldsymbol{B}$, with element

$$B_{i,j} = \begin{cases} 1 & , (i \to j) \in \texttt{\$ADMaps}, i \neq j \\ 0 & , \text{otherwise} \end{cases}$$

The element is one when a parent node i has an edge to a child node j in the tree; equivalently, it is one if there is an edge from vertex i to vertex j in the associated `$ADGraph` object constructed from the map. Otherwise, the binary transition is zero.

The, non-normalized, transition probability matrix, $\tilde{\boldsymbol{P}}$, is constructed using a

specified adaptation kernel, with probability density function f, with elements

$$\tilde{P}_{i,j} = \begin{cases} f(\boldsymbol{\alpha}_k|\mathcal{A}_i) & , b_{i,j} > 0, i \neq j \\ 0 & , \text{otherwise} \end{cases}$$

Here $\boldsymbol{\alpha}_k = \mathcal{A}_j \backslash \mathcal{A}_i$ (the complement of $\mathcal{A}_j$ with respect to $\mathcal{A}_i$) is the adaptation parameter of the nascent lineage in the transition from state i, with resident adaptation parameters $\mathcal{A}_i$, to state j with adaptation parameters $\mathcal{A}_j$. That is, $f(\boldsymbol{\alpha}_k|\mathcal{A}_i)$ is the probability that a nascent mutant lineage with adaptation parameters $\boldsymbol{\alpha}_k$ emerges given the resident phenotype composition $\mathcal{A}_i$.

Finally, the transition matrix is obtained by row-normalizing $\tilde{\boldsymbol{P}}$ so that

$$P_{i,j} = \begin{cases} \tilde{P}_{i,j} & , i \neq j \\ 1 - \sum_j \tilde{P}_{i,j} & , i = j \end{cases}$$

which yields a stochastic matrix. Note that the final normalization is correct because only non-trivial transitions have been retained in the adaptation map - that is, only those nascent lineages that survive invasion, and therefore modify the resident lineage composition, are retained, whereas it is implied that all other lineages emerge and become extinct without modifying the resident composition.

Ensembles

Constructing the stochastic process for a graph constructed from an `$ADEnsemble` object (Sections 3.4.1 and 3.4.3.2) is more straight-forward because the adaptation kernel is implicit in the contained transitions. Note, however, that this does mean that a new ensemble needs to be constructed for every adaptation kernel. After optimizing bin edges and constructing an `$ADGraph` object from the ensemble, the transition matrix is constructed by counting the number of edges so that $\boldsymbol{P}$ has element

$$P_{i,j} = \frac{N_{i,j}}{\sum_j N_{i,j}}$$

where $N_{i,j}$ is the number of edges from state i to state j, including the self-edges.

Note that the adaptation kernel required so that the Markov process constructed from an `$ADMap` is equivalent to that constructed from an `$ADEnsemble`, assuming the default (on-simplex uniform) kernel has been used for trajectory sampling, is

$$f(\boldsymbol{\alpha}_k|\mathcal{A}_i) = \sum_{\boldsymbol{\alpha}_l \in \mathcal{A}_i} P(\boldsymbol{\alpha}_k|\boldsymbol{\alpha}_l) P(\boldsymbol{\alpha}_l|\mathcal{A}_i)$$

Which is the probability that the resident with phenotype $\boldsymbol{\alpha}_l$, from the group of residents with phenotypes $\mathcal{A}_i$, is the progenitor of the mutant lineage multiplied by the probability that the mutant phenotype is $\boldsymbol{\alpha}_k$ given this progenitor.

The probability that a lineage is chosen from amongst the resident lineages to be the progenitor is equal to its fractional abundance in the population

$$P(\boldsymbol{\alpha}_l|\mathcal{A}_i) = \frac{\|\boldsymbol{x}_l\|}{\|\mathcal{X}\|}$$

and

$$P(\boldsymbol{\alpha}_k|\boldsymbol{\alpha}_l) = \frac{1}{|\mathcal{A}_{\tilde{S}}(\boldsymbol{\alpha}_l)|}$$

where $\mathcal{A}_{\tilde{S}}(\boldsymbol{\alpha}_l) \subset \tilde{\boldsymbol{S}}$ is the subset of adaptation parameters that are within `$maxNorm` Euclidean distance from $\boldsymbol{\alpha}_l$. Since these adaptation parameter vectors are equally likely to be drawn from the unit simplex using the (default) Dirichlet distribution we only need to take the reciprocal of the number of reachable vectors to calculate the probability that the nascent lineage's adaptation parameter vector is drawn.

Having constructed the transition matrix, the stochastic process has been defined (as a Markov chain object `$ADMarkov`) and standard methods [126] can be used to analyze adaptation.

3.4.4.1 Classification of states

The software includes methods for analyzing properties of states and classes (sets of states) in the Markov chain. The properties described below are important for calculating stationary distributions. Upon creation of the `$ADMarkov` object, states are inspected to determine accessibility and communication:

- A state j is accessible from i if there exists a path from i to j. In other words: if $P^n_{i,j} > 0$ for some $n \geq 0$. The notation $i \to j$ is used to show accessibility.

- States i and j communicate with each other if they are accessible from each other: $i \rightarrow j$ and $j \rightarrow i$.

The state space is then partitioned into disjoint communication classes, $C_k \subset \tilde{\boldsymbol{S}}$, which consist of sets of communicating states. If there is only one communication class, the Markov chain is irreducible; otherwise, it is reducible [126].

The Markov chain may contain periodic states. The period of state i is defined as

$$d(i) = gcd\{n \geq 1 : P_{i,i}^n > 0\}$$

where 'gcd' stands for the greatest common divisor. If $d(i) > 1$ the state is periodic; otherwise, the state is aperiodic. In other words, state i is periodic if all paths leading from i back to i are integer multiples of $d > 1$. Periodicity is a class property, so it suffices to calculate the period for one state in each communication class - all other states in that class will have the same period [126].

Some states in the Markov chain have the property that the chain will revisit those states with probability one (ie. an infinite number of times). These states are classified as recurrent [126]; if a state is not recurrent, then it is transient. State i is recurrent if the expected number of returns (R) to state i diverges[2]; conversely, i is transient if

$$E[R] = \sum_{n=0}^{\infty} P_{i,i}^n < \infty$$

Recurrence and transience are class properties so it again suffices to determine the status of one state of each communication class to determine the status of the entire communication class.

Note that every recurrent communication class is a closed subset of the state space; that is, it is impossible to reach any state outside of this closed subset from any state within the subset. This is fortunate since it effectively reduces the Markov chain to simpler, smaller, chains.

[2]Because we are working with finite state spaces the distinction between positive and null recurrence is not necessary.

State and class classifications are saved as properties of the `$ADMarkov` object and are important for calculating stationary distributions.

3.4.4.2 Stationary distributions

The state of the Markov chain evolves under the action of the transition matrix

$$\boldsymbol{v}_n^T \boldsymbol{P} = \boldsymbol{v}_{n+1}^T$$

where $\boldsymbol{v}_n$ is the (vertex) state distribution vector at time n. Let $\boldsymbol{\pi}$ be invariant under the action of the transition matrix

$$\boldsymbol{\pi}^T \boldsymbol{P} = \boldsymbol{\pi}^T$$

If, moreover, $\boldsymbol{\pi}$ satisfies $\sum_k \pi_k = 1$ then it is also a stationary probability distribution of the Markov chain. Software methods are included to determine the existence and uniqueness of stationary distributions, and, if they exist, to calculate their value(s). The classification of states and classes is used to control the algorithms for calculating stationary distributions. The different cases are discussed below.

Irreducible, aperiodic Markov chains have a unique stationary distribution. It is calculated as the left eigenvector, with associated eigenvalue equal to one, of the transition matrix. Standard methods for calculating right eigenvectors [127, 128] can be applied to calculate the left eigenvectors since

$$\begin{aligned} \boldsymbol{x}_L \boldsymbol{A} &= \lambda \boldsymbol{x}_L \\ (\boldsymbol{x}_L \boldsymbol{A})^T &= (\lambda \boldsymbol{x}_L)^T \\ \boldsymbol{A}^T \boldsymbol{x}_L^T &= \lambda \boldsymbol{x}_L^T \end{aligned}$$

And we can use $\boldsymbol{x}_R = \boldsymbol{\pi}$ and $\boldsymbol{A}^T = \boldsymbol{P}$.

For chains with a large number of states an alternative numerical method is also provided: the power iteration, which works well with sparse transition matrices. Iterative application of

$$\boldsymbol{v}_{n+1} = \frac{\boldsymbol{P}^T \boldsymbol{v}_n}{\|\boldsymbol{P}^T \boldsymbol{v}_n\|}$$

is guaranteed to converge

$$\lim_{n\to\infty} \boldsymbol{v}_n = \boldsymbol{\pi}$$

where $\boldsymbol{\pi}$ is the left eigenvector with associated eigenvalue equal to one.

The power iteration actually converges to the eigenvector associated with the dominant eigenvalue. However, in the case of irreducible Markov chains, the Perron-Frobenius theorem guarantees that 1 is an eigenvalue of $\boldsymbol{P}$ with multiplicity one [108]; moreover, all other eigenvalues will have an absolute value strictly less than 1. Therefore, the dominant eigenvalue will have a value of 1 and the power iteration converges to the stationary distribution.[3].

The power iteration method is numerically stable, however the speed of converge is determined by the magnitude of the second-largest eigenvalue. In practice, speed of convergence has not been an issue in applications encountered so far, but it is worth bearing in mind.

Irreducible Markov chains that are also periodic need to be treated a little bit more carefully. While there exists an invariant measure $\boldsymbol{\pi}$, which is unique up to multiplicative constants with $0 < \pi_k < \infty$ for all $k \in \tilde{\boldsymbol{S}}$ it may not always be a stationary distribution. Fortunately, for the finite Markov chains encountered in applications to adaptation dynamics, the invariant measure is always a stationary distribution as well (because recurrent states/classes in a finite Markov chain cannot be null recurrent). Note, however, that unlike aperiodic chains where $\lim_{n\to\infty} P^n_{i,j} = \pi_i$, the stationary distribution of a periodic chain does not give the long-term transition probabilities. Instead, the stationary distribution is interpreted as the average amount of time spent in each state and $\pi_k = 1/\mu_k$ where μ_k is the mean recurrent time of state k.

[3]The Perron-Frobenius theorem applies to stochastic matrices that have only strictly positive elements; however, transition matrices may contain some zero elements. The trick here is to apply the theorem to $\boldsymbol{P}^n$, which, for some $n \geq 1$, will have only strictly positive elements since every state in an irreducible transition matrix is reachable from every other state. Then, by the spectral mapping theorem, we can use the eigenvalues of $\boldsymbol{P}^n$ as the nth powers of the eigenvalues of $\boldsymbol{P}$. The eigenvectors are the same.

Therefore, irreducible Markov chains have the property that any initial (vertex) state distribution will converge to a unique stationary distribution.

Reducible Markov chains, on the other hand, need to be decomposed to their constituent chains for analysis. As before, let $\{C_k\} \subset \tilde{\boldsymbol{S}}$ be the set of disjoint communication classes. Let $r_1, r_2, \ldots$ be the recurrent communication classes and $t_1, t_2, \ldots$ the transient classes. The following results apply [126]:

1. Each recurrent communication class r_k has a stationary distribution $\boldsymbol{\pi}^{(k)}$ with support only on r_k when the Markov chain is restricted to r_k.

2. Each of the $\boldsymbol{\pi}^{(k)}$ is a stationary distribution of the full Markov chain.

3. Each linear combination $a_1\boldsymbol{\pi}^{(1)} + a_2\boldsymbol{\pi}^{(2)} + \ldots$ with $a_i \geq 0$ and $\sum_i a_i = 1$ is a stationary distribution of the full Markov chain.

4. All stationary distributions of the Markov chain are linear combinations of the 'restricted' distributions.

The methods for irreducible Markov chains can be applied to the restricted communication classes to calculate the family of stationary distributions. The limiting distribution, therefore, will always exist, but it will depend on the initial (vertex) state distribution vector. Note that in all cases the mass of the limiting distribution will be entirely in the recurrent communication classes and never in any of the transient classes [126].

It is possible to calculate exactly the limiting distribution for an arbitrary initial distribution vector, $\boldsymbol{v}$. Let $\alpha_k(i)$ be the probability that the chain, starting in state i, will eventually be absorbed in recurrent class r_k. Then the limiting transition probability from state i to state j is

$$\lim_{n\to\infty} P_{i,j}^n = \alpha_k(i)\pi_j^{(k)}$$

The jth component of this (limiting) distribution is given by

$$\pi_j = \sum_{i\in\tilde{\boldsymbol{S}}} \alpha_k(i) v_i \pi_j^{(k)}$$

That is, the contribution of the kth restricted stationary distribution is weighted by the probability of the chain starting in state i multiplied by the probability of state i being absorbed in class k.

Note that $\alpha_k(i) = 0$ if $i \in r_{k'} \neq r_k$ and conversely $\alpha_k(i) = 1$ if $i \in r_k = r_k$. Therefore calculating the $\alpha_k(i)$ is only difficult for the transient states. We can show that for any transient state t_i

$$\alpha_k(t_i) = p_{t_i,r_k} + \sum_{t_j} p_{t_i,t_j}\alpha_k(t_j)$$

This leads to a system of linear equations which can be solved to obtain the long-term absorption probabilities.

Software methods have been implemented to determine the existence and uniqueness of the stationary distribution of the Markov chain based on the classification of states and classes. If a unique stationary distribution exists then it is returned; otherwise, the family of stationary distributions is returned. When an initial (vertex) state distribution vector is provided the limiting distribution is calculated by either returning the unique stationary distribution or by calculating the linear combination of restricted distributions. Notable initial distribution vectors include the uniform distribution on the monomorphic vertices (ie. there must have been a *single* ancestral lineage, which is intuitive), and the initial distribution centered on a specific adaptation parameter vector of interest (usually an extremum or mid-point on the simplex) with probability one. Figure 3.4.9 summarizes the software's Markov process analysis procedure.

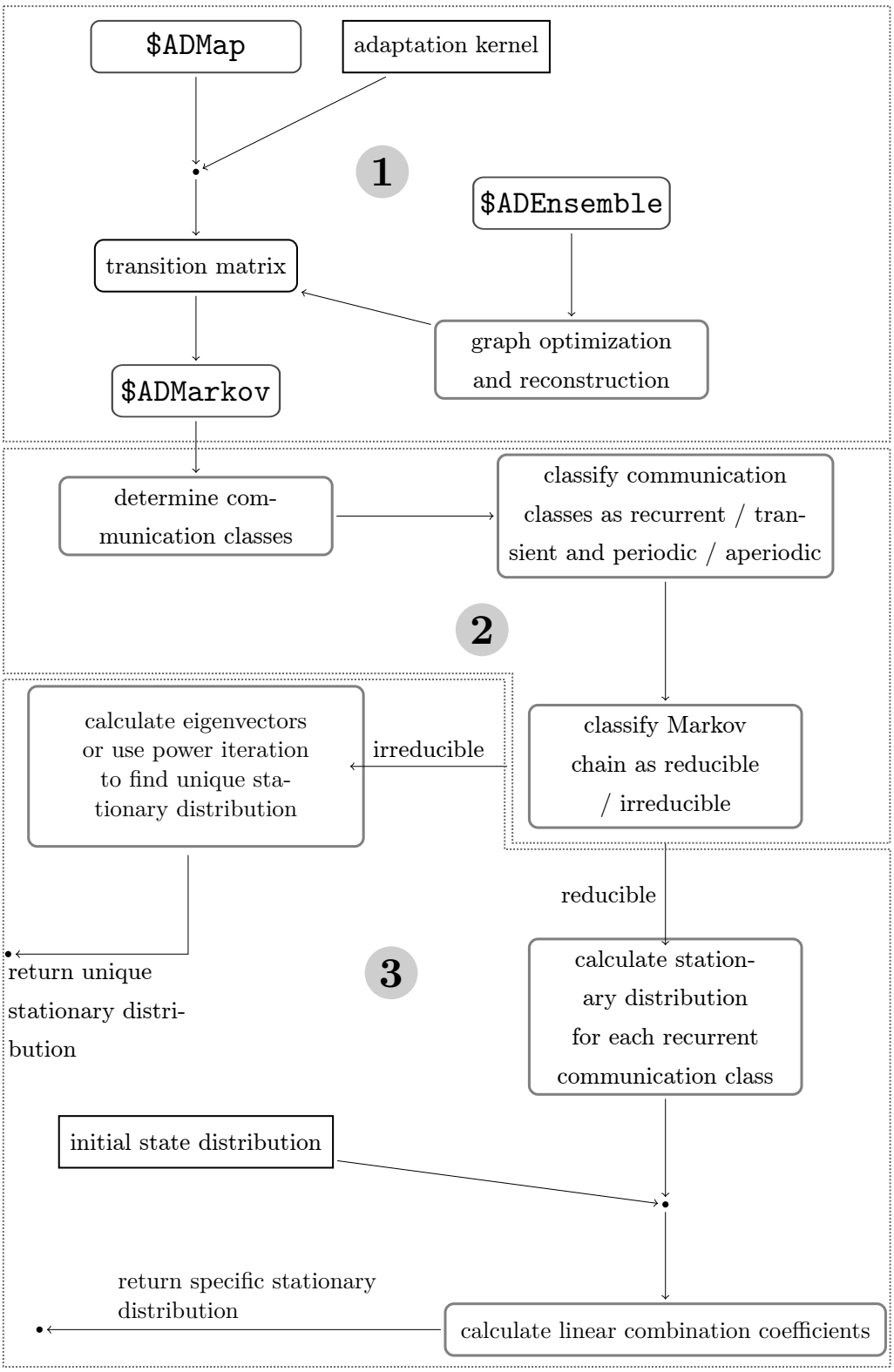


Figure 3.4.9: Summary flowchart for the Markov process analysis. *Continued on next page...*

1: Starting with an adaptation map (`$ADMap`), or an adaptation tree (`$ADTree`), we first construct the transition matrix between the dynamically-stable models. The transitions are weighted according to an adaptation kernel, which specifies the distribution of adaptation parameters around a progenitor lineage's phenotype. Starting with an ensemble of trajectories (`$ADEnsemble`), we generate an adaptation graph using the binning/optimization methods described in Section 3.4.3.2. The adaptation kernel has been pre-specified during the ensemble generation; subsequently, the transition matrix of this graph is constructed as described in Section 3.4.4. The discrete Markov process is now completely defined. **2:** The states of the process are partitioned into disjoint communication classes (Section 3.4.4.1), consisting of states that are mutually reachable from each other. Each class is then classified as stationary or transient, to indicate whether the adaptation process is 'trapped' in that class or if it will eventually escape; in addition, the periodicity of each class is determined, if the class is not aperiodic. **3:** If there is only one (recurrent) communication class then the process is irreducible and has a unique stationary distribution which can be calculated by solving the eigensystem (for example, using the power iteration3.4.4.2). If there are multiple (recurrent) communication classes, then there exists a family of stationary distributions, which is a linear combination of the stationary distribution of the constituent recurrent communication classes. The coefficients of the linear combination are determined from the probability that the Markov process becomes absorbed by each of the communication classes, given the initial state distribution. Once these are determined, the linear combination of stationary distributions gives the initial condition-specific distribution that we require.

3.4.5 Incorporating stochastic effects in graph construction

The graph construction methods described in Section 3.4.3 ignore some stochastic effects, such as accidental extinction, during the adaptation process. In this section, we will develop a more careful adaptation map parsing procedure that will correct the transition probability in the adaptation graph to account for: stochastic extinction probability, a possibly variable mutation rate, along with the mutation kernel that we have already described.

3.4.5.1 Deriving the transition probabilities

We begin by stating the corrected graph edge transition probability for a small unit of evolutionary time, dt. An edge may be added in the adaptation graph as a result of a rare mutant phenotype, r_y, emerging in a population of resident phenotypes, $r_{\boldsymbol{x}}$, which is assumed to be at steady state with lineage concentrations $\boldsymbol{x}^*$. The transition probability for that edge is:

$$P(r_y, dt|r_{\boldsymbol{x}}, \boldsymbol{x}^*) = P_I(r_y|r_{\boldsymbol{x}}, \boldsymbol{x}^*) \cdot P_s(r_y|r_{\boldsymbol{x}}, \boldsymbol{x}^*) \cdot \sum_{i=1}^{n} P_{m,i}(r_{x_i}, x_i^*) \cdot P'(r_y|r_{x_i}) \tag{3.6}$$

We begin with the term for the 'probability' of invasion, $P_I(r_y|r_{\boldsymbol{x}}, \boldsymbol{x}^*)$. This is the deterministic outcome of invasion contingent on the rare mutant having escaped stochastic extinction. In our approach, therefore, this is a binary term that can be obtained from the adaptation map, since it contains the set of all possible transitions/invasions. Accordingly, we set $P_I = 0$ if the mutant deterministically goes extinct, and $P_I = 1$ otherwise.

A rare mutant phenotype might be lost by accidental (stochastic) extinction due to an initially small number of individuals. The probability that the mutant phenotype will escape accidental extinction, $P_s(r_y|r_{\boldsymbol{x}}, \boldsymbol{x}^*)$, can be calculated with an application of the Moran process [6] in the limit of an infinite population sizes. In particular, for a single mutant phenotype, we have

$$P_s(r_y|r_{\boldsymbol{x}}, \boldsymbol{x}^*) = \lim_{N\to\infty} \frac{1 - r^{-1}}{1 - r^{-N}} = \frac{\lambda(r_y|r_{\boldsymbol{x}}, \boldsymbol{x}^*)}{\lambda_B(r_y|r_{\boldsymbol{x}}, \boldsymbol{x}^*)} \tag{3.7}$$

where we have used $\lambda(r_y|r_{\boldsymbol{x}}, \boldsymbol{x}^*) = \lambda_B(r_y|r_{\boldsymbol{x}}, \boldsymbol{x}^*) - \lambda_D(r_y|r_{\boldsymbol{x}}, \boldsymbol{x}^*)$ as the per-capita growth rate of the mutant, calculated from the difference of the per-capita birth, λ_B, and death, λ_D, rates. We evaluate the λ terms at the resident population's steady state and assume that they are constant throughout the stochastic process. Note that both the assumption of infinite population size and constant per-capita birth/death rates are approximations that need to be scrutinized more carefully. The infinite population size assumption makes sense in the context of our deterministic models; however, the constant birth/death rate is more troublesome – not least because these rates are non-linear functions of the nutrient variables, which change during the *later* stages of a successful invasion.

$P_{m,i}$ is the probability that the ith resident lineage generates a single mutant cell in a small interval of evolutionary time, dt. This probability is determined from a, possibly phenotype-dependent, mutation rate and the demographic details of the model at steady state:

$$P_{m,i}(r_{x_i}, x_i^*) = dt\mu(r_{x_i})\lambda_B(r_{x_i}, x_i^*) \cdot x_i^* \tag{3.8}$$

Here, $\mu(r_{x_i})$ is the mutation rate, which for full generality can be a function of the phenotype. $\lambda_B(p_{x_i}, x_i^*)$ is the per-capita birth rate of the ith resident, and x_i^* is its steady state concentration. We follow Dercole and Rinaldi [73] in deriving this probability, working with a demographic/evolutionary timescale separation and assuming that, to first order, the probability of obtaining more than mutation in a small interval is $\mathcal{O}(dt^2)$ and can be ignored in the infinitesimal limit of dt. Note that μ is the effective *phenotypic* mutation rate; however, see Section 3.4.5.3 for a discussion on adaptation on neutral genetic networks.

Finally, $P'(r_y|r_{x_i})$ is the probability that the mutant phenotype will be r_y, given that the parent lineage had phenotype r_{x_i}. This is the probability that we discussed earlier, given by the mutation kernel. To re-iterate, $P'(r_y|r_{x_i})$ is a family of distributions, D, parametrized by $\boldsymbol{d}$. Because we are working with a discretized phenotype space (the simplex $\tilde{S}$), we have to re-normalize the probabilities, so we have:

$$P'(r_y|r_{x_i}) = \frac{D(r_y|r_{x_i}; \boldsymbol{d})}{\sum_{y \in \tilde{S}} D(r_y|r_{x_i}; \boldsymbol{d})} \tag{3.9}$$

In earlier discussions, we used the uniform mixture kernel parametrized by the maximum mutation norm. In that case we had

$$D(r_y|r_{x_i};d_{\max}) = \begin{cases} c & ||r_y - r_{x_i}|| \leq d_{\max} \\ 0 & \text{otherwise} \end{cases} \tag{3.10}$$

where $d_{\max}$ is the range of the uniform kernel (maximum mutation norm), and c is a non-zero constant that cancels after re-normalization.

After applying the above corrections – to, essentially, the $P_I(r_y|r_{\boldsymbol{x}},\boldsymbol{x}^*)$ that are contained in the adaptation map – the edges of the adaptation graph can be interpreted as the transition probabilities for a small evolutionary interval. Importantly, unlike our original method for constructing graphs, the edge probabilities in the transition matrix, from vertex v_i to vertex v_j, are *not* row-normalized. The self-edge transition probabilities are then relatively large, and determine the characteristic time scale of adaptation.

With the transition probabilities from Equation 3.6 it is possible to create a discrete-state, continuous-time, Markov chain model of adaptation [129]. After multiplying the terms in Equation 3.6 we obtain the terms

$$P(r_y, dt|r_{\boldsymbol{x}},\boldsymbol{x}^*) = dt \cdot [\dots] \tag{3.11}$$

to first order dt. The infinitesimal generator matrix of the Markov chain is obtained from the matrix of transition probabilities, $\boldsymbol{P}$ (for interval dt), in the infinitesimal limit:

$$\boldsymbol{Q} = \lim_{dt\to 0^+} \frac{\boldsymbol{P}(dt) - \boldsymbol{I}}{dt} \tag{3.12}$$

where the probabilities are now taken over the phenotype composition vertices, $P_{i,j}$ for the transition implied when mutation r_y transforms the phenotype composition from that of vertex i to that of vertex j. Note that the dt terms in the limit will cancel since the self-edge probability is

$$P_{i,i} = 1 - \sum_j P_{i,j} = 1 - \sum_j dt \cdot [\dots] = 1 - dt \sum_j [\dots] \tag{3.13}$$

noting that the dt terms will cancel. This allows us to express a finite-state master equation as the forward Kolmogorov differential equations:

$$\frac{d\boldsymbol{P}(t)}{dt} = \boldsymbol{Q}\boldsymbol{P}(t) \tag{3.14}$$

which we can solve, given initial conditions. This interpretation permits us to make testable predictions about adapting systems, such as calculating the expected evolutionary time for an initial phenotype composition to reach the stationary vertex. However, these calculations are more difficult to do, requiring analysis of continuous-time Markov chains, and we have not pursued this course.

We note that our earlier approach for parsing the adaptation map (Section 3.4.4) amounts to generating the *embedded, discrete-state, discrete-time, Markov chain* of the continuous-time process described above [129]. Consequently, the state classification that we perform for determining the stationary vertex distribution in the discrete-time Markov process will hold for the continuous-time process as well. Importantly, it can be shown that the limit distribution of the (finite) embedded Markov chain is also a stationary distribution of the (finite) continuous-time Markov chain, provided that the embedded chain is irreducible.

3.4.5.2 Numerical methods

To include the above stochastic effects in our calculations, we must specify the per-capita birth and/or death rates. While we can numerically obtain $\lambda(r_y|r_{\boldsymbol{x}}, \boldsymbol{x}^*)$ from the simulation trajectory ($\lambda(t) = \dot{y}(t)/y(t)$) it is not possible to partition this number to λ_B and λ_D without the explicit form of the last two terms. One possibility is to require lineage rate equations in the general models from Equation 3.1 models to be specified in the following way:

$$\frac{d\mathcal{X}}{dt} = \mathcal{X}\left[\boldsymbol{\Lambda}_B(\mathcal{X}, \boldsymbol{\nu}, t; \boldsymbol{m}, \mathcal{A}) - \boldsymbol{\Lambda}_D(\mathcal{X}, \boldsymbol{\nu}, t; \boldsymbol{m}, \mathcal{A})\right] \tag{3.15}$$

With Λ_B and Λ_D being the expressions for per-capita birth and death rates that must be specified explicitly. With this separation of the terms, we can now numerically calculate all terms in Equation 3.6 from the invasion trajectories, allowing us to construct the continuous-time Markov chain representation. In particular, the methods that construct the adaptation map (Section 3.4.2) should

now also retain the $\lambda(r_y|r_{\boldsymbol{x}}, \boldsymbol{x}^*)$ terms, to be used when creating the adaptation graph.

An additional complication exists for the types of models we have looked at, stemming from population structure and the existence of multiple λ_i (one for each lineage variable in the structure). We have shown in other sections – Section **??**, Figure **??** and Section **??**, Figure **??** – that the invasion eigenvalues (λ_i) become equal on the invasion manifold, after an initial transient. The sign of the eigenvalue determines the success of invasion; however, the λ_i may initially have different signs. We discussed how evaluating the eigenvalues at/after their convergence gives the correct prediction for the invasion success. Similarly, in this case, the software must first determine the time point, $t^\dagger$, at which the $\lambda_i(t)$ have converged and use the corresponding $\lambda_{B,i}(t^\dagger)$ and $\lambda_{D,i}(t^\dagger)$ for the probability calculations. This could by achieved numerically by finding the time to convergence, $t^\dagger$:

$$t^\dagger = \min_t \left\{ \sum_{i \neq j} \frac{||\lambda_i(t) - \lambda_j(t)||}{\min\{|\lambda_i(t)|, |\lambda_j(t)|\}} \right\} \tag{3.16}$$

over the simulation trajectory for all pairs of per-capita growth rates. The denominator ensures that a trivial time point at steady state, where all $\lambda_i = 0$, is not returned.

3.4.5.3 Adaptation on neutral genetic networks

Our model of adaptation makes the simplifying assumption that evolution takes place in phenotype space. In reality, however, adaptation proceeds via evolution on degenerate neutral genetic networks [104, 130], as was demonstrated by the discovery of neutral networks in phenotype-genotype models of RNA secondary structure [131, 132]. Figure 3.4.10 shows a characteristic schematic of a neutral network wherein genotypes of equal (or marginally different) fitness are connected via single-nucleotide mutations. These neutral networks are, in turn, embedded in a wider network of phenotypes, each of which is characterized by a neutral network.

What are the implications of neutral genetic networks for phenotypic adaptation, particularly as they pertain to our modeling approach? Naively, we might

expect that exploration of large genetic networks would slow down the effective phenotypic mutation rate. However, Draghi *et al.* [130] demonstrated that the relationship between the size of a neutral network and adaptation time need not be monotonic. The probability that a mutation is neutral, q, is related to the size and connectivity of the neutral network. q can be thought of as the robustness of the phenotype to genetic mutation. The relationship between q and adaptation rate is determined by the ratio of phenotypes accessible from any single genotype in the focal network, K, compared to the number of all alternative phenotypes, P. For $K = P$, the relationship between robustness and adaptation rate is indeed monotonic; however, this condition implies that neutral mutations have no epistatic consequences, which does not agree with experimental evidence. On the other hand, for network topologies where $K < P$, the adaptation rate is a non-monotonic function of q, which is minimized for some intermediate value of robustness.

We can incorporate mutation on neutral genetic networks in our modeling approach by first re-interpreting $\mu(r_x)$ as the rate of genetic mutation, instead of the effective phenotypic mutation rate. The probability that a mutation is neutral can then be incorporated in the mutation kernel, which appears in Equation 3.9, where we can set

$$D(r_{y=x_i}|r_{x_i};\boldsymbol{d}) = q(r_{x_i}) \tag{3.17}$$

For full generality we can allow network robustness to vary between phenotypes; otherwise, we can have $q(r_{x_i}) = q$. The remaining $D(r_{y=x_i}|r_{x_i};\boldsymbol{d})$ of the mutation kernel can be taken as the expected probability of transition from a genotype in r_{x_i} to a genotype in r_y, with the expectation taken over the distribution of the genotype probabilities on the neutral network of r_{x_i}. When epistatic interactions are important (i.e. the case $K < P$ in *et al.* [130]) then, in the regime where neutral mutations are much more likely than non-neutral mutations, we could take the expectation over the limit distribution of the *isolated* neutral network of r_{x_i} (as in Van Nimwegen *et al.* [133]). When $K = P$, whereby every genotype in the local network can access the same phenotypic neighborhood, the mutant phenotype probabilities (relative to each other) do not depend on the distribution of the genotypes in the neutral network (but note that there is still a dependence between the neutral network size (robustness) and the ratio of neutral to non-

neutral mutation probability). The latter case is considerably easier to implement in our modeling framework.

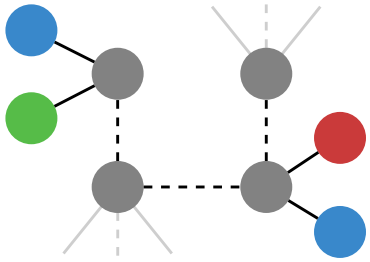


Figure 3.4.10: Schematic of a neutral genetic network. Circles represent genotypes and colors represent phenotypes. Genotypes of the same phenotype have the same fitness, or their fitness difference is small enough (relative to population size) to be effectively neutral. Genotypes on the same phenotype network are connected via a single nucleotide mutation (broken edges). Genotypes between phenotypes are also connected via single nucleotide mutations, but note that the phenotypic neighborhood of two genotypes on the same neutral network may be different.

3.4.5.4 Anticipated differences between the continuous- and discrete-time Markov processes

The transition probabilities, and resulting infinitesimal generator matrix, calculated from Equation 3.6 for the continuous-time Markov chain (CTMC) will be different from those calculated using the methods in Section 3.4.4 for the discrete-time Markov chain (DTMC). The two processes are related: our earlier method constructed the embedded Markov chain of the continuous process, and the two processes share many properties – particularly with regards to the classification and calculation of stationary vertex (state) probabilities.

In our DTMC formulation, we did not consider stochastic population effects and we ignored the mutation rate. Nevertheless, we can anticipate, as long as the $P_s(r_y|r_{\boldsymbol{x}}, \boldsymbol{x}^*)$ and $P_m(r_{\boldsymbol{x}}, \boldsymbol{x}^*)$ do not vanish for some $(r_y, r_{\boldsymbol{x}})$, that the long-term stationary behavior of graphs with a single stationary vertex (SSV) will not change. Indeed, the SSV graphs are irreducible and aperiodic DTMCs, and since there is only one absorbing state, the limit distribution of the CTMC will be the same as that of the DTMC. It is more difficult to anticipate how MSV graphs will be affected by the continuous formulation; however, we expect that the set of stationary vertices (vertices with a non-zero limit distribution) will persist, but

with a change in the actual probabilities at steady state.

In Chapter 4 we characterize the adaptation process in terms of paths through the Markov chain graph, and calculate the 'length' of the adaptation trajectory in units of invasion/mutation events. These results, obtained from the embedded DTMC of the continuous-time process, should not change in the CTMC formulation – in fact, constructing the embedded chain is how one would obtain the measures we calculated.

3.5 Summary of numerical methods

In this chapter we have shown how the process of phenotypic adaptation can be understood as a discrete Markov process whose states are dynamically-stable models. The states are each characterized by the phenotypic parameters of the extant lineages and they are connected via mutation/invasion events.

We have presented a set of numerical methods, implemented in a software package for the Wolfram Language that facilitate working with general models of adaptive dynamics. The methods allow flexible model specification, and automate model generation, simulation, resolution, and augmentation to generate adaptation maps and ensembles of adaptation trajectories.

In the last section we have shown how we use these data to generate the model's discrete Markov process. Methods were implemented for processes analysis - notably, state classification and stationary distribution calculation, under a variety of mutation/adaptation kernels and initial conditions.

In Chapter **??** we will apply the numerical methods to first compare with the theoretical results from Chapter **??** and then to analytically intractable models that support richer adaptation dynamics.

Chapter 4

Applications of numerical methods to models of adaptive evolution

4.1 Introduction

In this chapter we will apply the numerical methods described in Chapter **??** to three models of adaptive dynamics.

In Section 4.2 we will re-analyze the model with a single limiting nutrient and two alternative metabolic pathways that we first treated analytically in Section **??**. This serves as an introduction to the type and scope of analysis that the numerical methods allow us to perform. We will demonstrate how a more complete and comprehensive characterization of the adaptation process can be obtained by simulating the outcome of all possible invasion events (models) on a discretized parameter grid. We will visualize the adaptation process as a graph and then use discrete Markov process analysis to highlight how the mutation size, together with the ancestral (initial) phenotype distribution, determine the stationary outcome of adaptation. We will then conduct a parameter space exploration to investigate how model parameters affect adaptation.

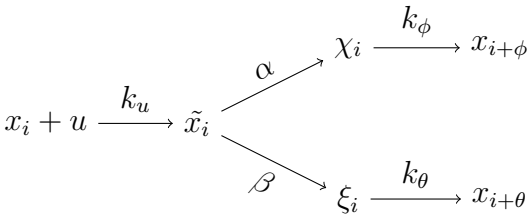
In Section 4.3 we will describe a new model whereby cells are growing on two nutrients. The adaptation parameter characterizes the 'preference' of a cell lineage to use each of the two nutrients. We begin by defining the model and highlighting

the analytical intractability to motivate our resort to numerical methods. We will demonstrate the application of numerical methods to an example parameter set, where we will show how lineage co-existence further complicates the characterization of the adaptation process. We will then randomly sample the parameter space and construct adaptation maps and graphs. We will then classify the graphs according to their stationary behavior and show how classes localize in parameter space.

Finally, in Section 4.4 we will extend the model with two nutrients up to three nutrients. We have adopted an approach similar to the that of the previous model (demonstration to build intuition, followed by parameter space sampling and graph classification) to characterize the scope and dynamics of adaptation. We will conclude by finding parameter subspaces that separate our graph classes.

4.2 Two alternative metabolic pathways

We now revisit the adaptation model with two alternative metabolic pathways that was first introduced in Section **??**. The model reactions are re-printed in Figure 4.2.1.



$$\chi_i \xrightarrow{k_\phi} x_1 + x_{i+\phi-n}$$

when $i + \phi > n$

$$\xi_i \xrightarrow{k_\theta} x_1 + x_{i+\theta-n}$$

when $i + \theta > n$

$$\emptyset \xrightarrow{p} u$$

$$\{x_i, \chi_i, \xi_i, u\} \xrightarrow{D} \emptyset$$

Figure 4.2.1: Schematic for the structured model with alternative metabolic pathways showing reaction channels. There are two alternative metabolic pathways that we will name the *phi* (ϕ) and the *theta* (θ) pathway. Successful metabolism of one unit of nutrient along the phi pathway forces the cell to transition from state i to state $i + \phi$. We refer to ϕ as the yield of the, similarly named, phi pathway. The same holds for the theta pathway, with yield θ. If the state transition exceeds some critical state, n, then a daughter cell is produced and the mother cell transitions to the state corresponding to the remainder of the "state mass" after replication.

Recall that the adaptation (phenotype) parameter is

$$s_x = \frac{\alpha}{\alpha + \beta} \tag{4.1}$$

and can be thought of as the metabolic pathway 'preference'. s_x may be interpreted as the success rate of a Bernoulli trial, where success means that the next unit of nutrient is metabolized using the phi pathway (χ_i variables) and failure means that it is metabolized using the theta pathway (θ_i variables).

4.2.1 Demonstration of numerical methods

A parameter set (Table 4.1) has be chosen to demonstrate the application of the numerical methods presented in Chapter **??**. We have selected simple parameters, emphasizing a rate/yield trade-off in the metabolic pathways themselves: the 'phi' pathway has a low yield ($\phi = 1$) but a relatively high rate ($k_\phi = 10$) while the 'theta' pathway has a higher yield ($\theta = 2$) and a lower rate ($k_\theta = 1$). We know from Section **??** that similar combinations of parameters are more likely to have adaptively stable strategies that are 'generalists' - i.e. not $s_x = 0$ or $s_x = 1$, where the cells make exclusive use of a single pathway.

symbol	parameter	value
ϕ	ϕ pathway yield	1
θ	θ pathway yield	2
k_ϕ	ϕ pathway metabolic rate	10
k_θ	θ pathway metabolic rate	1
k_u	nutrient binding rate	1
p	nutrient concentration	1
D	dilution rate	0.08

Table 4.1: Parameters for the demonstration model.

Our analysis from Section **??** has shown that it is possible to derive an expression for the invasion fitness of pathway-preference phenotype s_x entirely in terms of the nutrient steady state in a single-resident environment; to be precise, a nascent mutant lineage with phenotype s_y will invade, and drive the resident lineage with phenotype s_x to extinction if

$$u_x^* > u_y^* \tag{4.2}$$

where u_x^* and u_y^* are, respectively, the nutrient steady states in the resident-only and mutant-only chemostat models; therefore, the nutrient steady state as a function of the s_x parameter completely determines the possible adaptation dynamics of the model (Figure 4.2.2).

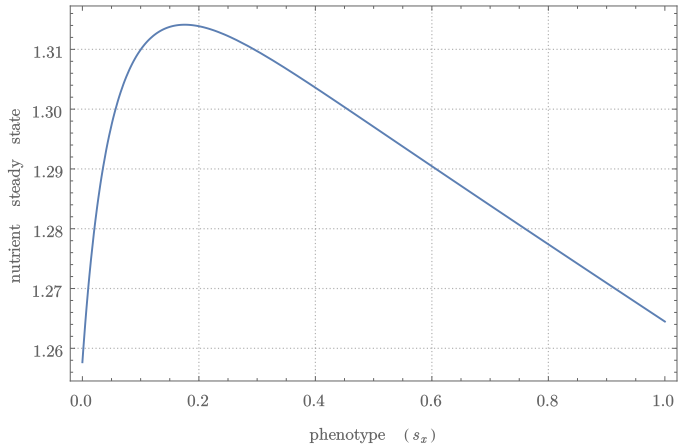


Figure 4.2.2: Nutrient steady state for resident-only models with resident phenotype s_x. A nascent mutant lineage with lower nutrient steady state, compared to its immediate ancestor, will invade and drive the original resident to extinction; conversely, if the mutant has a higher nutrient steady state it will itself become extinct.

The two minima of the model lie at the interval extremes, $s_x = 0$ and $s_x = 1$, and are both surrounded by larger nutrient steady state values for almost the entire unit line. Notice, however, that $s_y = 0$ can invade $s_x = 1$, because it is the global minimum; therefore, the phenotype that preference the θ pathway entirely ($s_x = 0$) is the expected long-term outcome of the adaptation process.

However, the properties of the adaptation process, most importantly the magnitude of the maximum allowed difference between the mutant phenotype and the ancestral phenotype, restricts the observable adaptation dynamics. This is intuitively obvious in light of Figure 4.2.2: the adaptation process can escape local minima provided it can make large enough 'jumps'.

The adaptation process can be visualized by plotting the trajectory of the resident in phenotype space (Figure 4.2.3). Sampling adaptation trajectories provides an intuition for the possible adaptation dynamics but not a complete, or quantitative, picture of the entire space of dynamics.

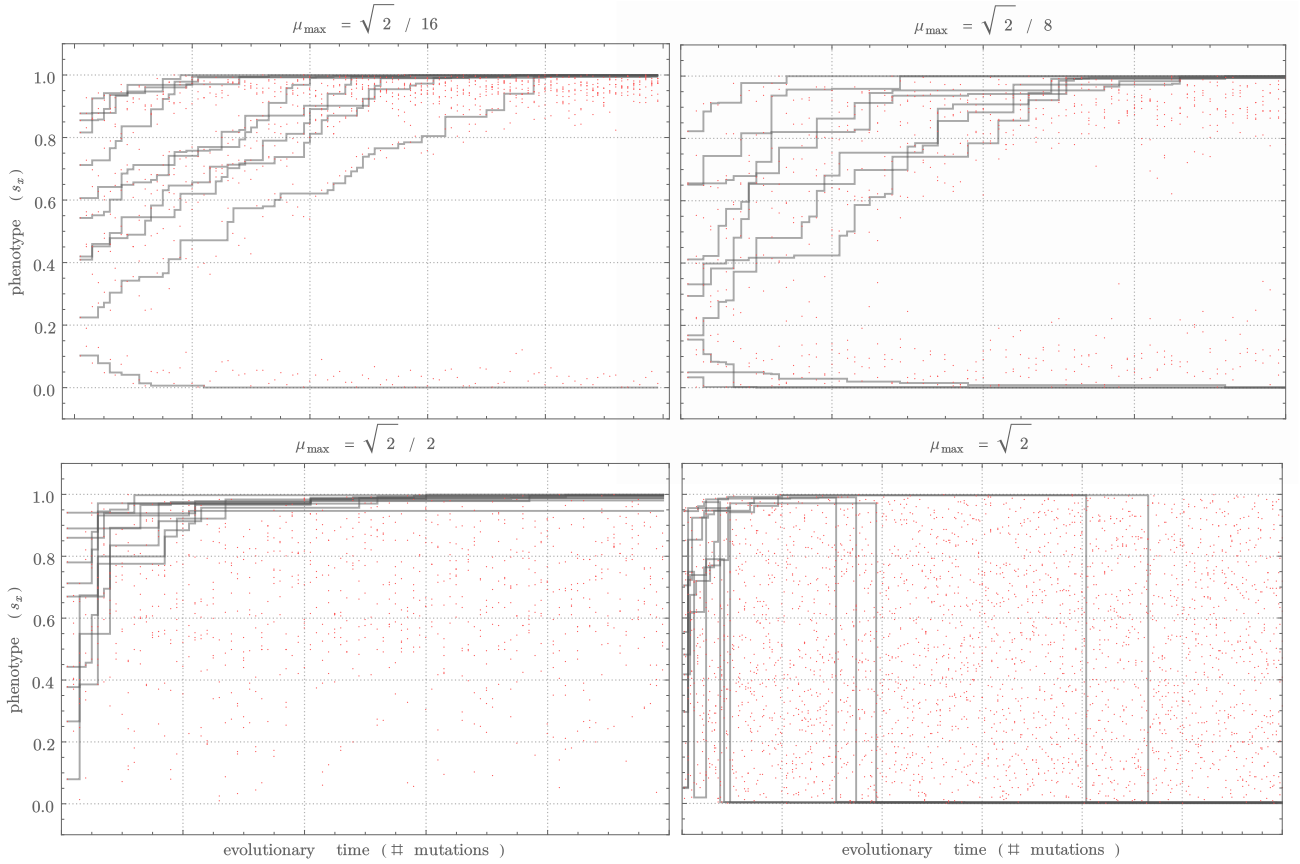


Figure 4.2.3: Sampled trajectories in the demonstration model using a uniform adaptation kernel. Each panel shows ten independent trajectories (black lines) in phenotype space. The title of each panel shows the maximum mutation norm, μ_{max}, which is maximum allowed Euclidean distance between the mutant and progenitor phenotype on the unit simplex, up to a maximum of $\sqrt{2}$. Because only a single resident can be supported at any given time, the ten independent trajectories are plotted together. The red points are the phenotypes of the nascent mutant lineages observed during the sampling process. For small values of μ_{max} the trajectories converge to either $s_x = 0$ or $s_x = 1$. Larger maximum mutation norms eventually lead to a single stationary state for all trajectories at $s_x = 0$.

A more complete view of the adaptation process can be obtained by discretizing the phenotype space and inspecting all possible (invasion) models using the discretized grid. The adaptation map for the demonstration parameter set was constructed using the dynamic programming algorithm described in Section **??** and the process graphs plotted for a few values of the maximum mutation norm (Figure 4.2.4). The adaptation graphs show the possible paths that the resident phenotype may follow, starting - in this case - from any value on the discretized grid, before it is absorbed by one of the stationary vertices.

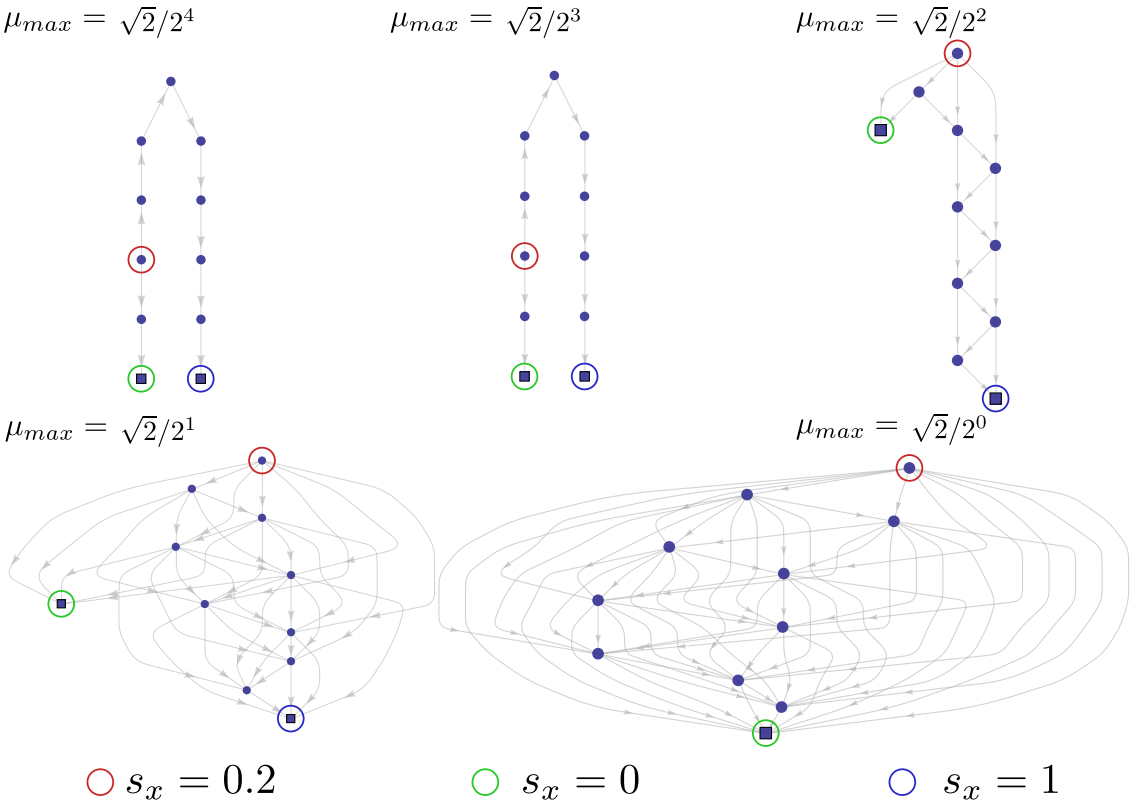


Figure 4.2.4: Graphs constructed from the adaptation map for the demonstration parameter set. The maximum mutation norm (μ_{max}) permitted is shown above each graph. Each vertex represents a single resident phenotype. A vertex receives an edge from another vertex if it is possible, via a single mutation, to change phenotypic composition of the population from that of the first vertex to that of the second vertex. Square vertices are stationary vertices in the discrete Markov process with a uniform initial distribution (any vertex may be the starting point of adaptation). The value of the μ_{max} determines which vertices are stationary.

While the graphs show the stationary vertices for a given maximum mutation norm, and the possible paths that the process can realize to these vertices, a more careful investigation reveals that the stationary vertex probabilities can vary

considerably depending on the actual value of μ_{max} as well as on the assumptions we make about the distribution of the ancestral phenotype (Figure 4.2.5).

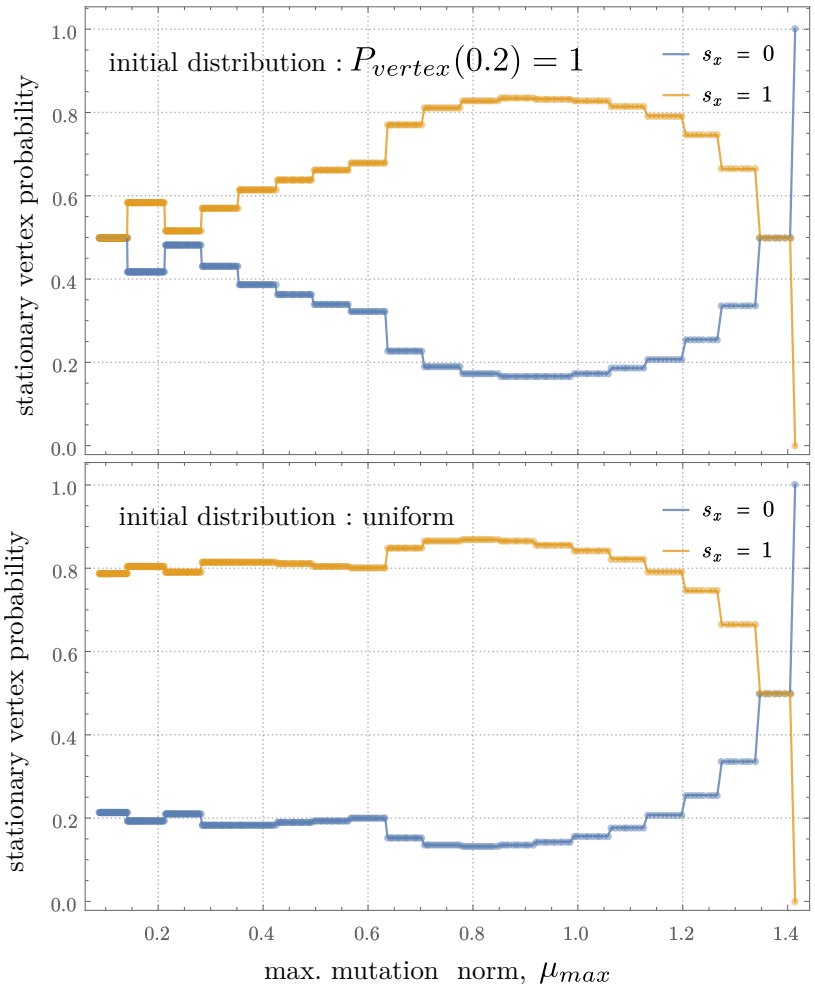


Figure 4.2.5: Stationary probabilities for the two possible stationary vertices ($s_x = 0$ and $s_x = 1$) as a function of the maximum mutation norm for two initial vertex distributions. The top panel shows the probabilities for the adaptation process that has an initial distribution centered on the $s_x = 0.2$ vertex (red highlight in Figure 4.2.4). The bottom panel shows the probabilities for the process where every vertex has an equal probability of being the ancestral (initial) one. The initial distribution is not important for the largest possible maximum possible mutation norm ($\mu_{max} = \sqrt{2}$) since every mutant phenotype is reachable from every other phenotype on the grid.

In conclusion, the full suite of numerical methods reveals more information about the quantitative aspects of adaptation that would not otherwise have been easily available using only the analysis and naive numerics from Section **??**.

4.2.2 Parameter space sampling

The procedures from the previous subsection were applied to 5,000 random parameter sets in order to investigate the dependence of various aspects of the adaptation process on model parameters.

The adaptation maps were used to construct the graphs of the adaptation process, for chosen values of the maximum mutation norm, μ_{max}. For simplicity, the stationary vertex distribution was calculated in each case using the uniform initial distribution. Figure 4.2.6 shows the average graphs of the adaptation process compiled from the ensemble of 5,000 individual graphs.

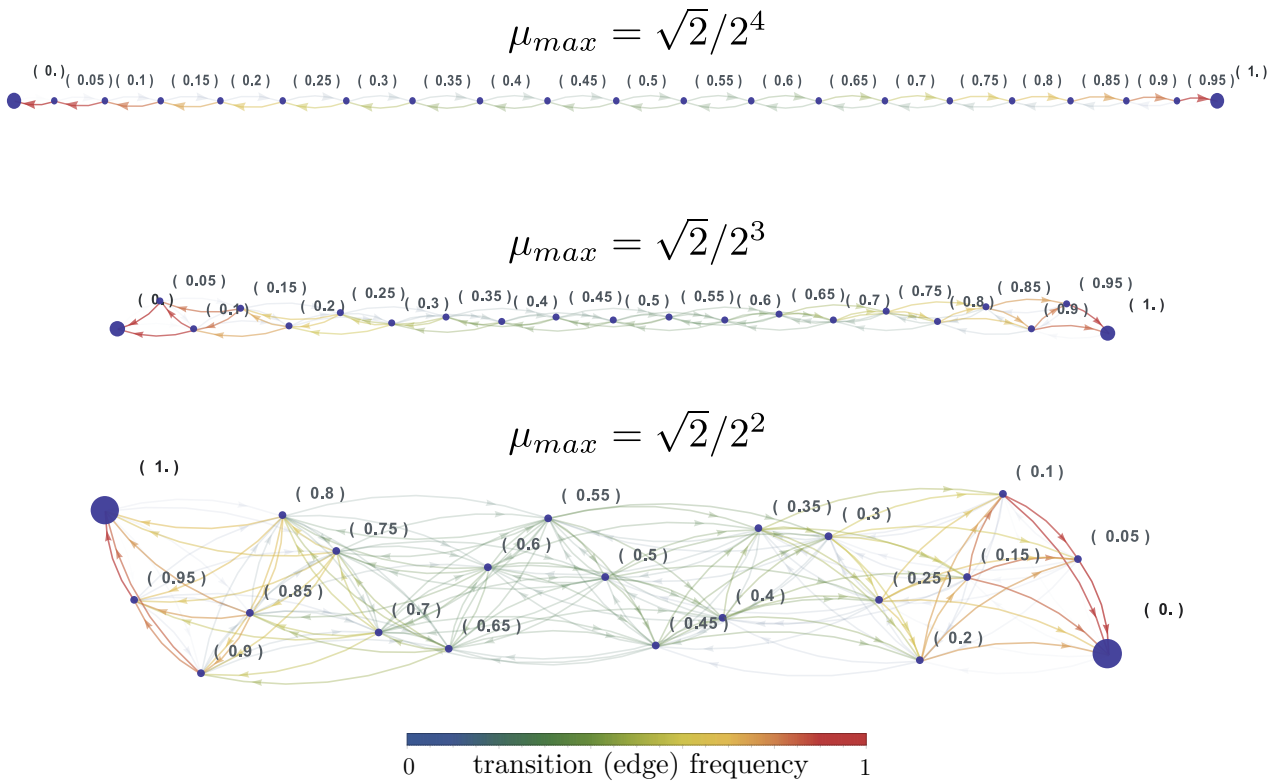
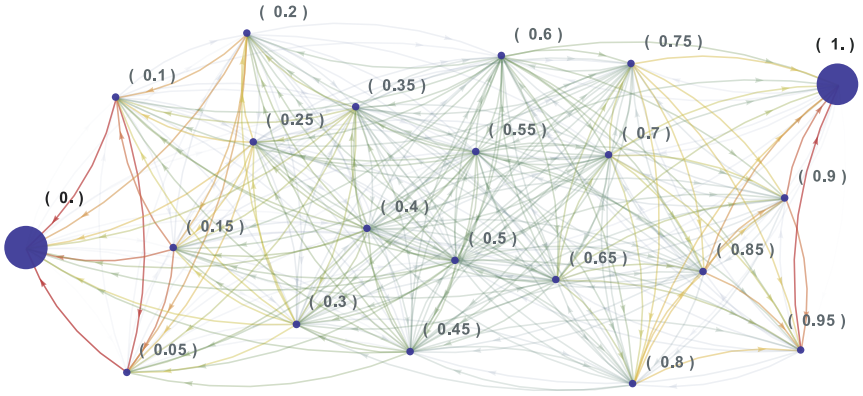
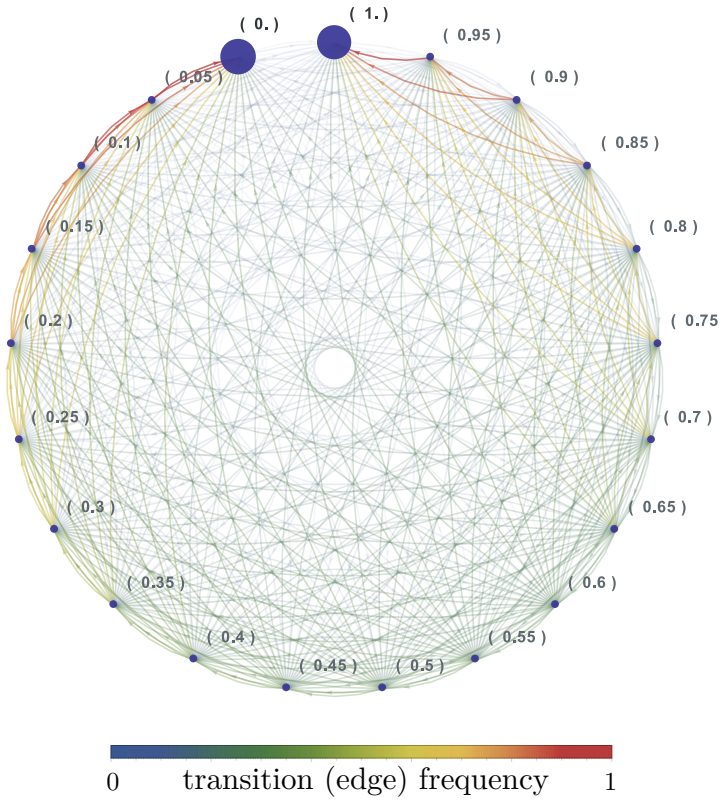


Figure 4.2.6: Average graphs constructed from the ensemble of graphs produced from the 5,000 adaptation maps. The maximum mutation norm, μ_{max}, is shown above each average graph. Vertex size is proportional to the average per-vertex stationary probability across the ensemble. Edges are colored according to the observed frequency in the ensemble - i.e. red (1) if the edge is present in all graphs. The vertices for the two 'specialist' phenotypes ($s_x = 0$ and $s_x = 1$) that have the highest per-vertex stationary probability. *Figure continues overleaf...*

$$\mu_{max} = \sqrt{2}/2^1$$



$$\mu_{max} = \sqrt{2}/2^0$$



The two most frequently observed stationary vertices, accounting for their stationary probabilities, were the ones corresponding to the two 'specialist' phenotypes ($s_x = 0$ and $s_x = 1$). From the average graphs of the map ensemble we may conclude that for a randomly chosen parameter set and an observed phenotype s_x the adaptation process will, on average, transition to a phenotype $s_y < s_x$ if s_x is closer to the $0-$edge and, conversely, transition to $s_y > s_x$ if s_x is closer to the $1-$edge. In other words, assuming the model parameters are not known, there is more uncertainty with regard to predicting the future state of the adaptation process when a phenotype of intermediate value is observed compared to observed phenotypes that lie closer to the 'specialist' extremes.

Most adaptation maps in the ensemble produced graphs with a single stationary vertex - almost always $s_x = 0$ or $s_x = 1$. About 3% of the maps had graphs that, for at least some of the tested maximum mutation norms, supported multiple stationary vertices - again, usually, both $s_x = 0$ and $s_x = 0$. Figure 4.2.7 shows the location of these graphs in parameter space. Almost all of the maps that had graphs with multiple stationary vertices had model parameters where one of the two pathways had a higher yield, but a lower metabolic rate, relative to the other pathway. In those cases, the stationary vertex probability varied considerably depending on the initial distribution used and on the μ_{max} value. The demonstration parameter set selected for the previous subsection was in the 3% that supported multiple stationary vertices and Figure 4.2.4 shows a typical progression of the adaptation graphs produced by parsing the model's adaptation map with different values of μ_{max}, while Figure 4.2.5 shows how unpredictably the stationary vertex probabilities can change. However, no general trend for these parameter sets was detected with regard to the preference of one type of pathway over the other when comparing large or small μ_{max}.

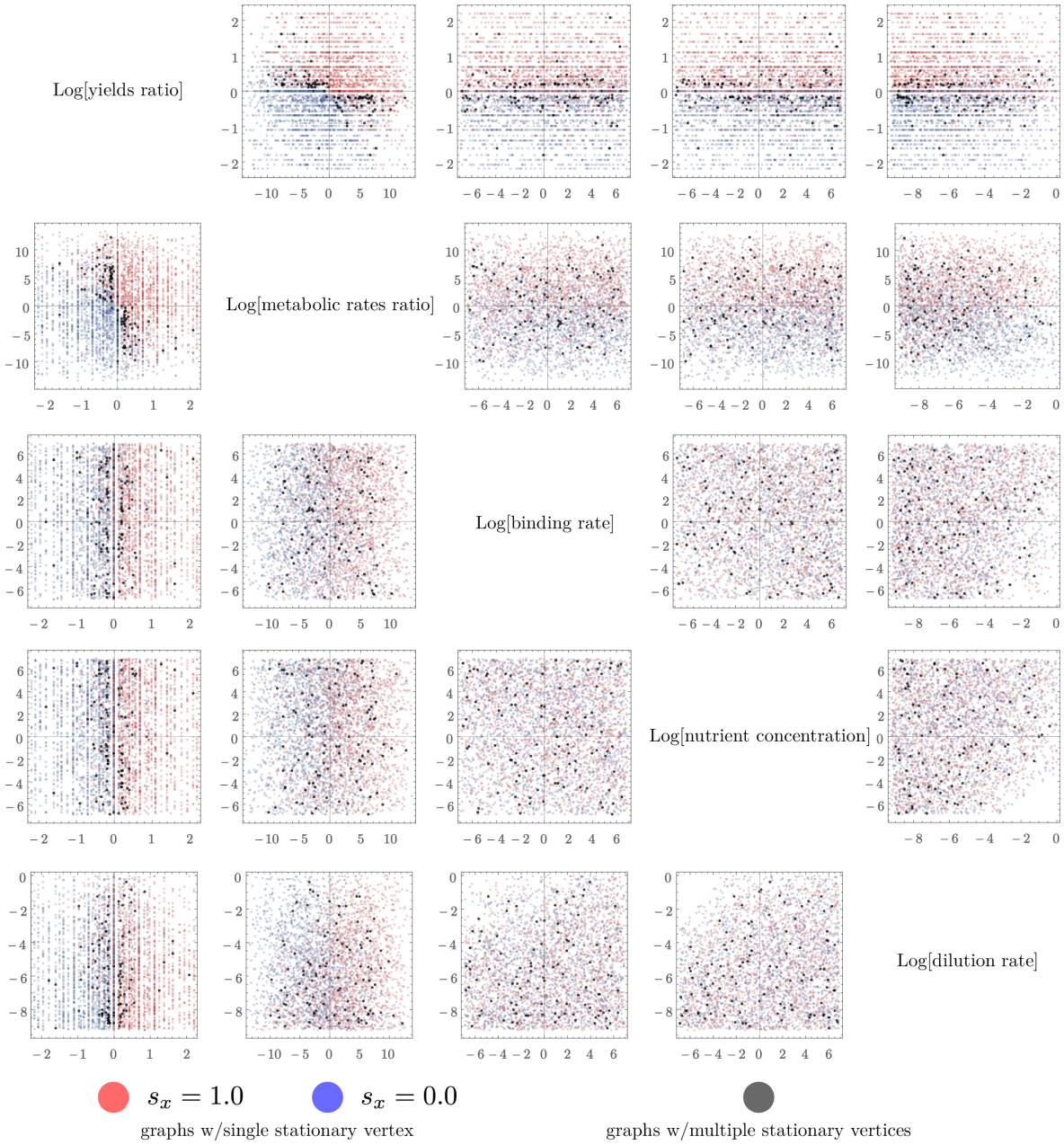


Figure 4.2.7: Scatter plot matrix of the ensemble maps' model parameters. The graphs were constructed using $\mu_{max} = \sqrt{2}/2^3$ and the stationary vertex distribution was calculated using a uniform initial distribution. Almost all of the stationary vertices for the graphs with a single stationary vertex were at the phenotype space boundaries (blue ($s_x = 0$) and red ($s_x = 1$) points). The black points show the maps that had multiple stationary vertices for at least one value of μ_{max}.

4.2.3 Section summary and discussion

In this section we have revisited the adaptive dynamics model with cells growing on a single limiting nutrient but two alternative metabolic pathways.

The analytical results from Section **??** provided an expression for the invasion condition between any nascent mutant phenotype and any resident phenotype. However, the invasion condition does not describe the adaptation process, particularly when mutations are large. As we have demonstrated, the mutation kernel - in our case, the maximum mutation norm of a Dirichlet kernel - determines the stationary vertex (phenotype) probabilities (Figure 4.2.4). Moreover, the initial (ancestral) vertex distribution plays a similarly critical role in determining the stationary outcomes of the adaptation process (Figure 4.2.5).

The average graphs (Figure 4.2.6) constructed from parameter space sampling give an overall impression of the adaptation process. We found that most model parameters give rise to adaptation processes that have metabolic pathway specialists as the adaptively stable strategies. We also found that when each metabolic pathway maximizes either yield or rate, but not both at the same time, that the graphs were more likely to support multiple stationary vertices (Figure 4.2.7). In those cases, the maximum mutation norm and initial vertex distribution are even more important for determining the stationary outcomes.

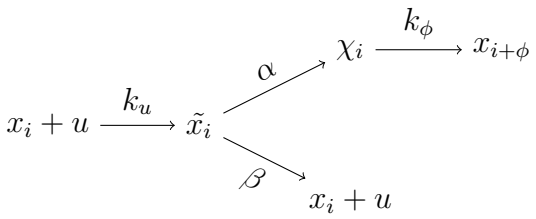
4.3 Two substitutable nutrients

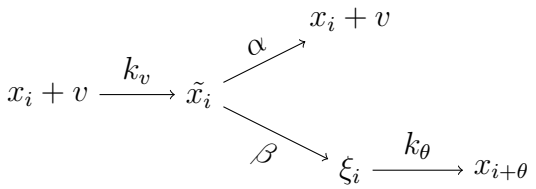
So far we have looked at models of adaptation where the phenotype determined how a cell behaved after binding and internalizing a nutrient. In this section we will consider a model where there are two types of nutrients in the environment and the phenotype determines how the cell interacts with these nutrients before they are internalized.

This model has many similarities with the models of two alternative metabolic pathways: the nutrients [u and v] are characterized by a metabolic yield [ϕ and θ] and a metabolic rate [k_ϕ and k_θ]. In addition, we allow for different nutrient binding (encounter) rates [k_u and k_v], whereby the nutrient molecules bind receptors on the cell surface and (may) become internalized; moreover, the two nutrients are added to the chemostat at different nutrient influx concentrations [p_u and p_v]. Finally, as before, we have a constant dilution rate [D] that washes out all model components from the chemostat.

Each lineage is characterized by a (unique) phenotype that determines the probability that a cell will metabolize a unit of nutrient that it has just encountered. Intuitively, this can be thought of as the fraction of cell surface receptors that bind and transport the nutrient in question. We will assume, for simplicity, that the receptors can bind and transport only one type of nutrient. This phenotype can be parametrized on the unit interval: $s_x \in [0, 1]$, with $s_x = 0$ indicating that all receptors bind the v nutrient, and $s_x = 1$ indicating that all receptors bind the u nutrient.

The reaction channels are shown schematically in Figure 4.3.1 and a cartoon of the model is shown in Figure 4.3.2.





$$\chi_i \xrightarrow{k_\phi} x_1 + x_{i+\phi-n}$$

when $i + \phi > n$

$$\xi_i \xrightarrow{k_\theta} x_1 + x_{i+\theta-n}$$

when $i + \theta > n$

$$\emptyset \xrightarrow{p_u} u$$

$$\emptyset \xrightarrow{p_v} v$$

$$\{x_i, \chi_i, \xi_i, u, v\} \xrightarrow{D} \emptyset$$

Figure 4.3.1: Schematic for the model with two alternative nutrients showing reaction channels. The chemostat contains two types of nutrients (u and v). The cell encounters the u, v nutrients at a rate proportional to their concentration and binding rate. When one unit of nutrient, u, for example, is encountered, it will be internalized a fraction $s_x = \frac{\alpha}{\alpha+\beta}$ of the times, equal to the fraction of cell surface receptors that are specialized for u; otherwise, the nutrient unbinds and is returned to the unbound pool of u nutrients in the chemostat. As with the previous models, successful metabolism of the nutrient results in progression of the cell along the growth state space by a number of states equal to the yield of the nutrient (ϕ for u and θ for v). When the cell transitions to a state that exceeds some growth threshold, n, the cell divides, which results in one daughter cell in the basal state (x_1) and a mother cell with the 'leftover' growth state mass. We have again assumed that the intermediate species, $\tilde{x}_i$, is at quasi steady state with respect to the rest of the system and acts as a decision point where the lineage phenotype acts.

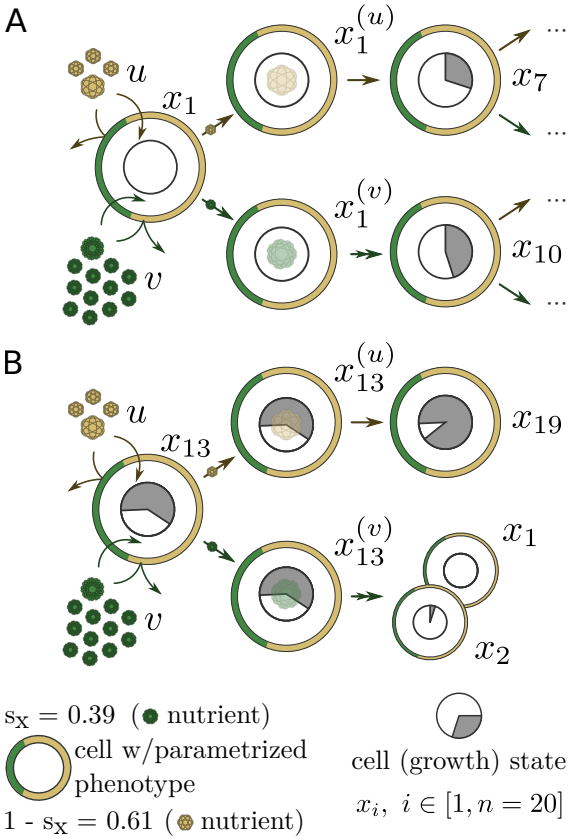


Figure 4.3.2: Cartoon of model with two substitutable nutrients. The two types of nutrients (u – yellow, and v – green) are internalized after binding the cell's receptors with a probability equal to the phenotype's value (s_x for u and $1 - s_x$ for v), which corresponds to the fraction of nutrient transporters dedicated to one of the two nutrients. ***A***. After a cell has taken up a nutrient, it will metabolize it at a nutrient-specific rate and receive a nutrient-specific yield. In this example, we have set the nutrient yields to 6 for the u nutrient (yellow) and 9 for the v nutrient (green). ***B***. Cells traverse the growth state space, shown here for a maximum of 20 states, and divide after they cross the 20th state. For example, a cell in state 13 will divide to two cells if it metabolizes a v (green) nutrient (since $13 + 9 > 20$) but not if it metabolizes a u (yellow) nutrient. Now shown in this cartoon are the constant nutrient influx rates and the constant dilution rate.

The equations for this model are shown below.

$$\begin{aligned}
\dot{\boldsymbol{x}} &= -\boldsymbol{x}(uk_u s_x + vk_v(1-s_x) + D) + k_\phi \boldsymbol{\Phi}\boldsymbol{\chi} + k_\theta \boldsymbol{\Theta}\boldsymbol{\xi} \\
\dot{\boldsymbol{\chi}} &= \boldsymbol{x}(uk_u s_x) - \boldsymbol{\chi}(k_\phi + D) \\
\dot{\boldsymbol{\xi}} &= \boldsymbol{x}(vk_v(1-s_x)) - \boldsymbol{\xi}(k_\theta + D) \\
\dot{u} &= p_u - Du - uk_u s_x \sum_i^n x_i \\
\dot{v} &= p_v - Dv - vk_v(1-s_x) \sum_i^n x_i
\end{aligned} \tag{4.3}$$

$\boldsymbol{\Phi}$ and $\boldsymbol{\Theta}$ are the $\boldsymbol{\chi}$- and $\boldsymbol{\xi}$-specific transition matrices respectively:

$$\boldsymbol{\phi}_{i,j} = \begin{cases} 1 & : j+\phi = i \\ 0 & : \text{otherwise} \end{cases} + \begin{cases} 1 & : j+\phi-n = i \\ 0 & : \text{otherwise} \end{cases} + \begin{cases} 1 & : j+\phi > n \quad \&\& \quad i = 1 \\ 0 & : \text{otherwise} \end{cases} \tag{4.4}$$

$$\boldsymbol{\theta}_{i,j} = \begin{cases} 1 & : j+\theta = i \\ 0 & : \text{otherwise} \end{cases} + \begin{cases} 1 & : j+\theta-n = i \\ 0 & : \text{otherwise} \end{cases} + \begin{cases} 1 & : j+\theta > n \quad \&\& \quad i = 1 \\ 0 & : \text{otherwise} \end{cases} \tag{4.5}$$

which implement the modulo arithmetic of state progression and replication - as with the two alternative pathways model.

We have again made use of the quasi steady state approximation [110] on the intermediate cellular species, $\tilde{x}_i$, to implement 'decision' points in growth state space progression (similar to Section **??**). The nutrient preference/specialization phenotype is parametrized to

$$s_x = \frac{\alpha}{\alpha+\beta} \tag{4.6}$$

,to express the fraction of cell's surface receptors that bind the u nutrient, and to

$$1 - s_x = \frac{\beta}{\alpha+\beta} \tag{4.7}$$

to express the fraction that bind the v nutrient. We will use $s_y, s_z, \ldots$ to denote the phenotypes of additional lineages. The implicit assumption of the quasi

steady state approximation is that the 'acceptance' or 'rejection' of the nutrient by the cell surface receptors is instantaneous, or otherwise assumed into the k_u, k_v parameters.

The two nutrients in our model are (perfectly) substitutable, using the terminology of Leon and Tumpson [134], which we take to imply that the cells can grow using either of the two nutrients. Indeed, the relationship between growth and nutrients falls on a spectrum: on one end, nutrients are perfectly substitutable; on the other, both nutrients are essential and are both required for growth [135]. In this section we will only deal with the former case of substitutable nutrients.

In Chapter 2 we derived the invasion condition for models with a single limiting nutrient (Equations 2.70, 2.99). In particular, we found that a nascent mutant lineage will invade - and drive the resident lineage to extinction - if it has a positive initial growth rate in an environment conditioned by the resident lineage. The invasion condition was cast in terms of nutrient steady states from 'monoculture' (only one lineage) models:

$$u^{*resident*} > u^{*mutant*}$$

In other words, the mutant will invade if it conditions the environment (nutrient u) to a lower steady state value ($u^{*mutant*}$), compared to the resident ($u^{*resident*}$), when mutant and resident phenotypes are grown separately. This allowed us to make a prediction about the outcome of competition/invasion from prior measurements [46, 47], and thus describe the scope of adaptation dynamics.

We also demonstrated that an invading mutant lineage quickly converges to an invariant invasion manifold ($\mathcal{M}_I$, Equation 2.58) which is equivalent to the mutant manifold ($\mathcal{M}_M$, Equation 2.60). In other words, the distribution of the mutant population among the cell states during invasion is the same as the distribution of the mutant population when growing alone. Because we could analytically calculate the invariant manifold we could reduce the state-structured models to simpler models and derive invasion conditions.

We attempted to derive an analytical expression for the invasion condition for this model following the methods from Chapter 2. In principle, invasion is successful when the per-capita growth rates are positive during at least the early part of the invasion - i.e.

$$\frac{\dot{y}_i(t)}{y_i(t)}\Big|_{u=u^{*resident*},v=v^{*resident*}} > 0 \tag{4.8}$$

during which the nutrient variables (u, v) are close to the steady state values of the resident model.

Similarly to earlier models, we found that the mutant lineage quickly approaches an invariant invasion manifold (Equation **??**). Unlike the single-nutrient models, however, the invasion manifold is *not* equivalent to the mutant manifold and it exists only transiently (Figure 4.3.3).

The mutant per-capita growth rates converge on the (transient) invasion manifold - i.e.

$$\frac{\dot{y}_i(t)}{y_i(t)}\Big|_{\mathcal{M}_I} = \lambda, \forall i \tag{4.9}$$

according to numerical evaluations (Figure 4.3.3, second row). When λ is positive, the mutant initially grows and invasion is successful; conversely, when λ is negative, the mutant fails to invade.

We interpret the above as follows. An initial resident lineage grows to steady state and 'conditions' the environmental variables (nutrients) according to its phenotype (nutrient use parameter, s_x). A mutant lineage, with phenotype, s_y, appears at an initially relatively low concentration. While it is scarce relative to the resident lineage, it will approach an invariant cell state distribution, which depends on the 'fixed' values of $u = u^{*resident*}$ and $v = v^{*resident*}$. This is the invariant invasion manifold. If the mutant per-capita growth rates on this manifold are positive, then the mutant will grow to a significant fraction of the resident and will start appreciably changing the nutrient concentrations. The invasion manifold will then begin to transition to the mutant's own manifold and cell state distribution. If the per-capita growth rates are negative, then the invading lineage will become extinct (and will never reach its own manifold).

Because of the invasion/mutant manifold inequivalence, we could not derive an analytical expression for the invasion condition, since calculating the invasion manifold was not possible - not least because calculating expressions for $u^{*resident*}$ and $v^{*resident*}$ was not feasible. For this reason, we turned to numerical methods for generating and analyzing adaptive dynamics.

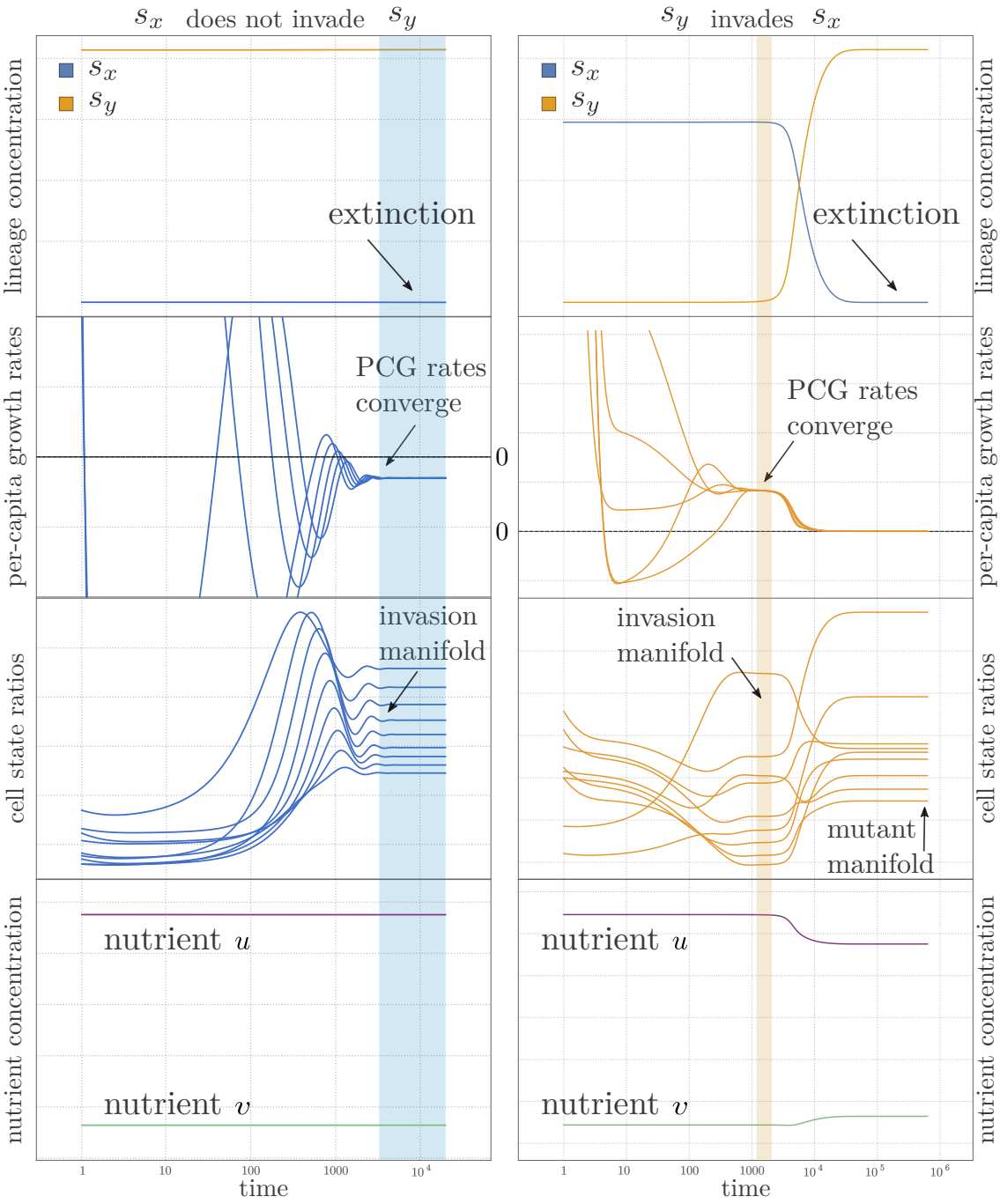


Figure 4.3.3: The invariant invasion manifold is transient and not equivalent to the mutant manifold. Simulated system trajectories for example parameters (Table 4.2), and two example phenotypes ($s_x = 0.9, s_y = 0.4$). With these parameters and phenotypes, s_y can invade a resident s_x, but s_x cannot invade a resident s_y. The per-capita growth (PCG) rates ($\dot{x}_i/x_i$ and $\dot{y}_i/y_i$) converge on a transient invariant manifold, where the cell state distribution ratios are constant(shaded regions). When the (converged) PCG rates are positive, invasion is successful (right panel); conversely, when these are negative, invasion fails.

4.3.1 Numerical analysis of an example parameter set

We will first restrict the application of the numerical methods software to a single parameter set (Table 4.2). We will demonstrate the various components of the software and their utility over the standard methods of adaptive dynamics analysis.

symbol	parameter	value
p_u	u concentration	3.8×10^{-3}
p_v	v concentration	1.0×10^{-2}
ϕ	u yield	1
θ	v yield	7
k_u	u binding rate	1.3×10^{2}
k_v	v binding rate	2.4×10^{-3}
k_ϕ	u metabolic rate	8.0
k_θ	v metabolic rate	1.3
D	dilution rate	6.1×10^{-4}

Table 4.2: Parameters for the two-nutrient demonstration model. In this parameter set, the u nutrient appears in the chemostat at a concentration (p_u) that is one order of magnitude lower than that of the v nutrient (p_v). The u nutrient provides a lower yield ($\phi = 1$ growth state per metabolic reaction) compared to v ($\theta = 7$ states per reaction). On the other hand, the u nutrient encounters and binds the cell's surface receptors much faster ($k_u = 1.3 \times 10^2$) compared to the v nutrient ($k_v = 2.4 \times 10^{-3}$), and it is also metabolized faster once internalized ($k_\phi = 8.0$ compared to $k_\theta = 1.3$. The choice of parameters for the demonstration example, therefore, does not permit a trivial prediction of which nutrient will be preferred.

Unlike models with a single nutrient, where a nascent mutant lineage with phenotype s_y would invade a resident phenotype s_x if it was better able to deplete the nutrient ($u_y^* < u_x^*$, Chapter **??**), the steady-state nutrient concentrations of isolated populations in this model are not informative (Figure 4.3.4). Figure 4.3.5 illustrates this point by comparing the actual outcome of pairwise invasions to the nutrient steady state concentrations from isolated resident models (Figure 4.3.4). Notably, the point $s_x \approx 0.3$, which appears to be adaptively stable (non-invasible) on the pairwise invasibility plot in Figure 4.3.5 does not correspond to a nutrient minimum, or any simple combination of nutrient values, according to Figure 4.3.4.

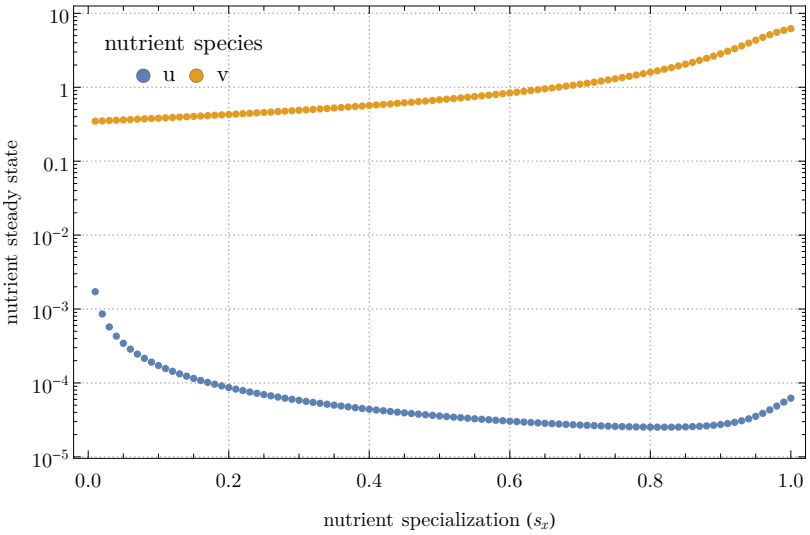


Figure 4.3.4: u (blue) and v (orange) nutrient concentration steady states of isolated resident model simulations as a function of the nutrient use phenotype. To generate these curves, single resident lineages, with phenotypes s_x, were simulated to steady state and the resulting nutrient steady state concentrations were plotted.

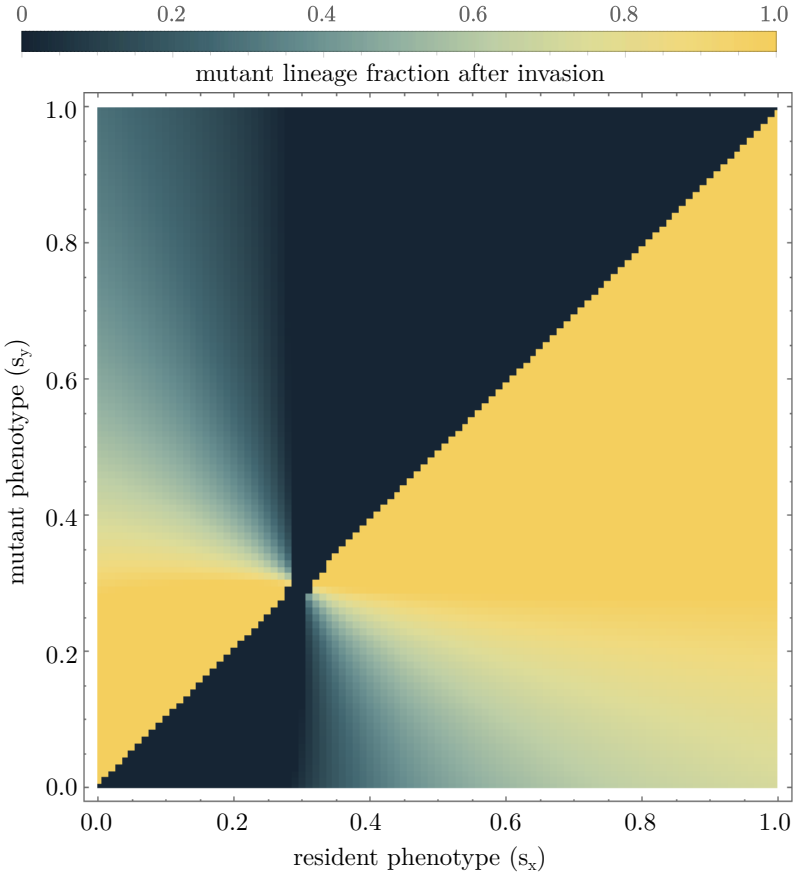


Figure 4.3.5: Pairwise invasion plot between a nascent mutant lineage (phenotype s_y) and an established resident lineage (phenotype s_x). Note the regions of co-existence where both lineages are maintained in the population.

Ecological co-existence, such as the one seen in Figure 4.3.5, has been a long-running focus of mathematical ecology [40, 42, 45]. In resource-consumer formulations of population models, co-existence conditions for populations growing on multiple nutrients are still being sought - some with notable, but still not universal, applicability [136, 137] for a subset of simpler models and parameter ranges. Indeed, much of the difficulty stems from the transience of the invasion manifold and its inequivalence to the mutant manifold. In our own case, we face the added difficulty from having structured consumer (population) models, which makes calculating nutrient steady states more difficult. The accepted, but not proven convention, is that co-existence between two lineages is possible if they are each mutually able to invade the other when rare [73]. In our numerical approach, however, we will not make this assumption; instead, we will rely only on empirical evidence of co-existence from model simulations.

Because of the presence, and prevalence, of co-existence between mutant and resident lineages, the pairwise invasion plot (Figure 4.3.5) does not adequately describe the dynamics and long-term state of adaptation: crucially, it does not contain any information about the fate of a nascent mutant lineage in an environment with two or more co-resident lineages. Therefore, an alternative method for presenting and summarizing the adaptation process is required for models where co-existence is possible.

We can gain some intuition about the model's adaptive dynamics by sampling random trajectories (Figure 4.3.6) using the methods described in Chapter 2 (section **??**). When the maximum allowed mutation step is small, adaptation takes the resident phenotype to approximately $s_x = 0.3$ via a series of monomorphic populations (only one resident lineage). A diversification event takes place at this point and two co-residents, with phenotypes in the neighborhood of $s = 0.3$, maintaining a stable and non-invasible coexistence. For larger maximum mutation steps, we observe early emergence of dimorphism and eventual establishment of a co-residency between phenotypes $s_x = 0.4, s_y = 0.0$. Therefore, the size of the maximum mutation norm (step) has an important effect on the dynamics, as well as the steady state, of the adaptation process.

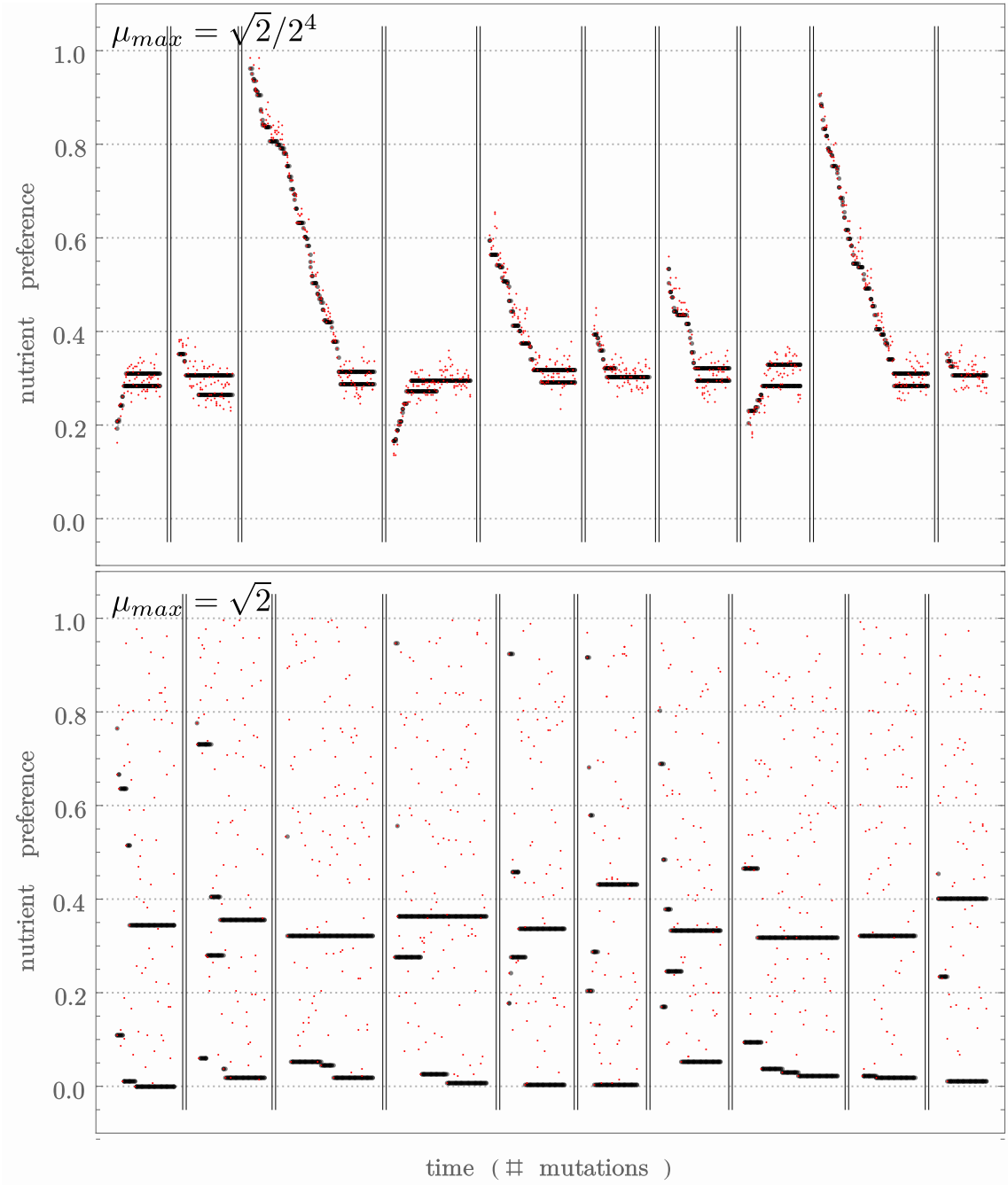


Figure 4.3.6: Sampling random trajectories in the demonstration model for a large (*top*) and a small (*bottom*) maximum mutation norm. Ten random trajectories are shown, delimited by double vertical lines, which have been started from random ancestral lineages and simulated to (approximate) adaptive steady state. Black dots show the established resident lineage(s). Red dots show the nascent mutant lineages. A random mutant lineage is generated at the beginning of every iteration whose phenotype is drawn from the uniform distribution on the unit simplex centered on the ancestral lineage. When multiple co-residents are present in the population, one is first chosen randomly with probability equal to its fractional abundance in the population to be the progenitor of the next nascent mutant.

To fully characterize the adaptation process we have to describe the outcome of all possible invasion events. However, as discussed in Chapter 2, this is prohibited in the absence of analytical expressions for the invasion fitness because of the continuous phenotype space; a compromise, then, is to discretize the phenotype space, which in this case is the unit line, and assume that phenotypes are constrained on the grid elements. Figure 4.3.7 shows the graphs of the adaptation map constructed with the dynamic programming algorithm (Chapter **??**, section **??**) using the demonstration parameter set for this model. In this example, the adaptation map was parsed to produce a graph using a uniform (on the simplex) mutation kernel (Section **??**), and various values for the maximum mutation norm. After parsing and graph construction, the vertices are classified (Section **??**) and the stationary vertex distribution is calculated - assuming a uniform initial vertex distribution on the monomorphic vertices - i.e. on the vertices containing a single resident lineage, which implies that the adaptation process must start from a single ancestral lineage. In interpreting the graphs it is important to note that while all possible invasion events are depicted (within our maximum mutation norm) not all parts of the graphs may be reachable because of the limitations imposed by the initial vertex distribution. This can be seen in Figure 4.3.7 (first panel, $\mu_{max} = \sqrt{2}/16$) where the dimorphic (red) vertices cannot be reached by the adaptation process that starts from a monomorphic (blue) vertex.

When the maximum mutation norm is small ($\leq \sqrt{2}/2^3$) the graph of adaptation dynamics appears highly structured. A monomorphic approach to the $s_x = 0.30$ vertex marks the early dynamics of adaptation and implies that the path of adaptation is highly predictable. Moreover, assuming the ancestral (initial) vertex is itself monomorphic, the $s_x = 0.30$ vertex at the end of the monomorphic 'chain' precludes the co-existence of more than one lineage at a time since $s_x = 0.30$ is a stationary vertex that will trap all trajectories. For maximum mutation norms larger than $\sqrt{2}/2^3$, however, the early monomorphic vertices can access the dimorphic vertices since these are no longer 'protected' by the $s_x = 0.30$ vertex; in fact, this vertex is no longer stationary. Of the dimorphic vertices, then, a Markov process analysis shows that only $\{s_x = 0.40, s_y = 0.00\}$ is a stationary population - it cannot be invaded by any mutant lineage.

Producing and comparing graphs from adaptation maps at different maximum mutation norms facilitates a qualitative understanding of the selective pressures

that drive adaptation in our model. In this case, for example, it can be shown that two phenotypes can stably coexist - one of which is a specialist in using the v nutrient exclusively ($s_y = 0.0$), while the other phenotype ($s_x = 0.40$) has as a small preference for the v nutrient but is also capable of using the u nutrient. It is not, however, possible for two lineages to coexist when their nutrient preferences are closely-matched because of the resulting conflict for the same niche. Only when mutations can alter the metabolic preference of a cell sufficiently far from those of the parental lineage can the adaptation process enter the dimorphic regime; subsequently, there is only one possible long-term adaptive outcome as the two phenotypes come to rest in a non-invasible configuration.

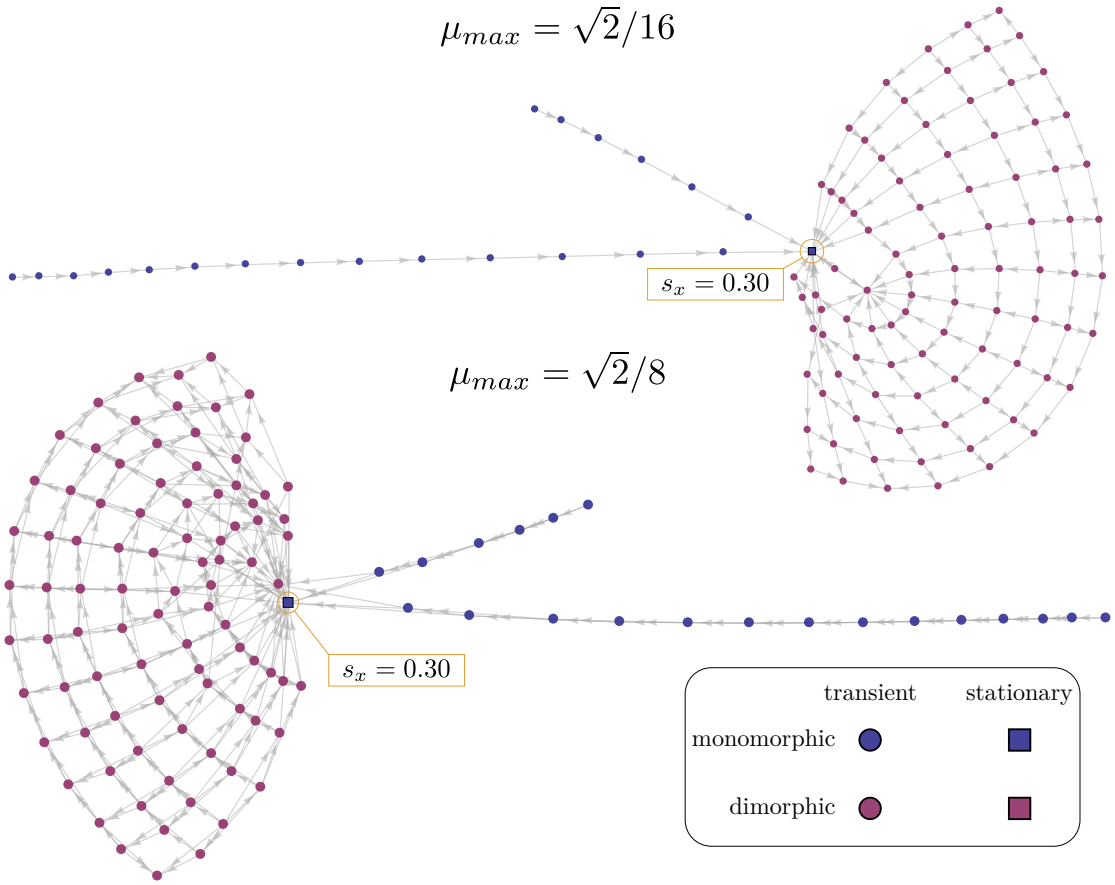
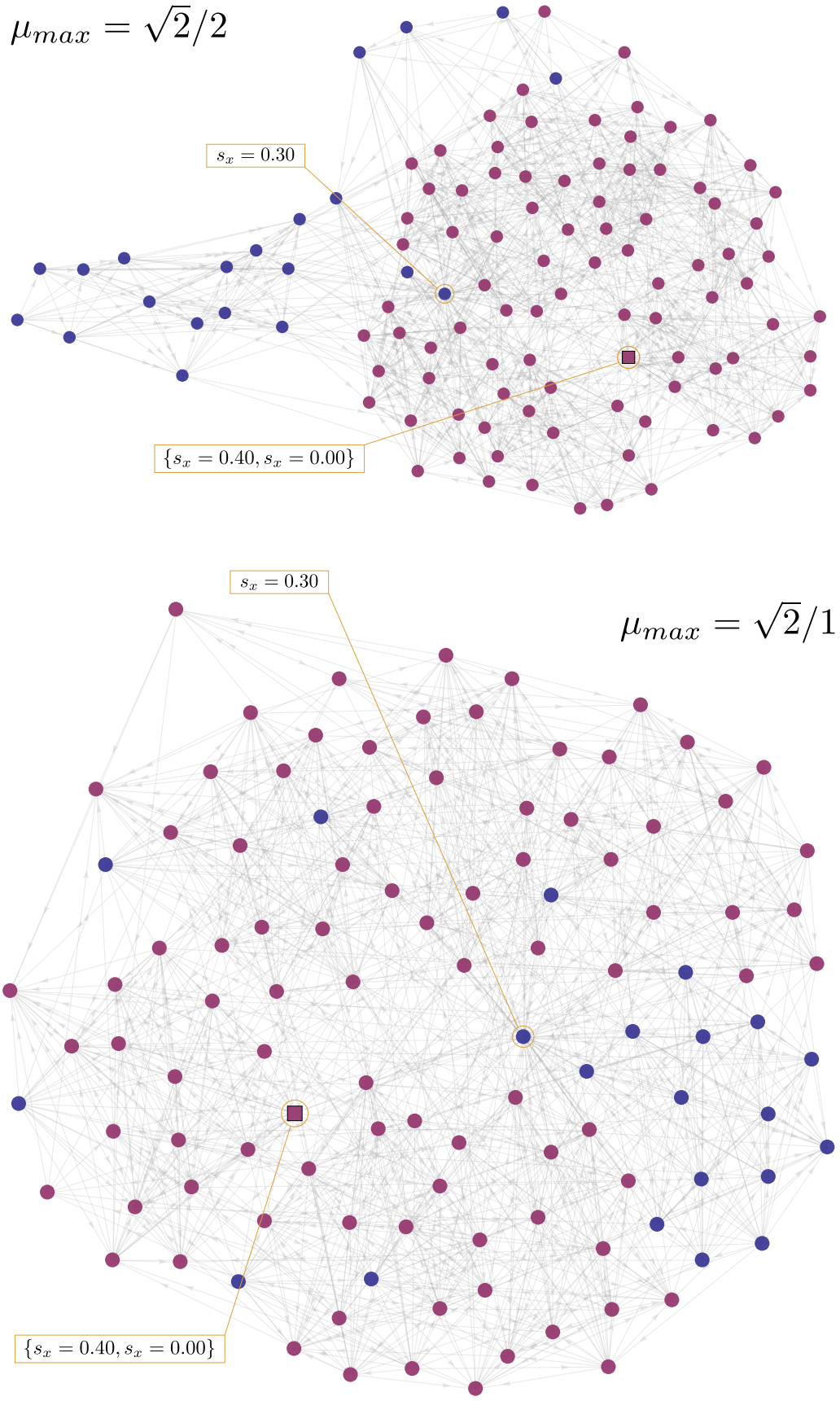


Figure 4.3.7: Graphs of the adaptation map constructed using the demonstration parameter set for the two alternative nutrients model. For each graph μ_{max} indicates the maximum permissible mutation norm (Euclidean distance in phenotype space.). Vertices represent dynamically stable models, colored according to the co-residency level (blue for single-resident (monomorphic) steady states and purple for two-resident (dimorphic) steady states). A vertex accepts an edge from another vertex if a nascent mutant lineage in the source vertex can change the steady state lineage composition to that of the destination vertex. Using an initial vertex distribution that is uniform only on the monomorphic (blue) vertices, the stationary vertex distribution for the discrete Markov process was calculated and the stationary vertices have been rendered as squares. Visually, stationary vertices have no outgoing edges. The phenotype space (unit line) was discretized using a 0.05 interval, giving 21 possible monomorphic vertices and $\frac{21^2-21}{2} = 220$ dimorphic vertices. *Figure continues overleaf*


$\mu_{max} = \sqrt{2}/2$
$s_x = 0.30$
$\{s_x = 0.40, s_x = 0.00\}$
$s_x = 0.30$
$\mu_{max} = \sqrt{2}/1$
$\{s_x = 0.40, s_x = 0.00\}$

4.3.2 Parameter space sampling and global analysis

To investigate the range of adaptive dynamics in the model of two alternative nutrients 5000 parameter sets were randomly sampled and their adaptation maps were computed on a discretized simplex. The graphs were computed at a range of maximum mutation norms and the stationary vertex distributions were calculated assuming a uniform initial distribution on the monomorphic vertices. We used a discretization grid with 10 points on the unit line. Taking advantage of our software's parallel-processing capabilities, the calculations required approximately 15 hours, running on two 2.8GHz quad-core Xeon E5462 processors (approximately 3×10^6 models simulated).

To summarize the findings visually, Figure 4.3.9 shows average graphs constructed from the 5000 individual graphs for four maximum mutation norms. A graph was produced from each adaptation map, using the uniform (on the simplex) mutation kernel and a maximum mutation norm, μ_{max}. The graph's stationary vertex distribution was calculated assuming the monomorphic uniform vertex distribution. We counted the number of graphs in which a particular edge appeared (transition from one dynamically stable model to another) for the space of all possible edges. The average graph color-codes the edge frequency (fraction of graphs in which the edge appears) to show which transitions are more likely to be observed for a random parameter set. We also calculated the mean per-vertex stationary distribution by taking the average of the stationary probabilities for a vertex across the $5,000$ graphs. When the vertex was not stationary, or when the vertex did not appear at all in the graph, we used a 0 stationary probability (for that graph). Each vertex's size in the average graph is proportional to its mean stationary probability.

We used the uniform mixture kernel to parse each adaptation map (resulting from a randomly sampled parameter set) with 11 maximum mutation norm values (setting the range of the support of each uniform distribution in the mixture) distributed logarithmically in $[\sqrt{2}/8, \sqrt{2}]$. For each of the $5000 \times 11 = 55000$ graphs, we calculated the stationary vertex distribution and hierarchically labeled the long-term stationary behavior, as shown in Figure 4.3.8.

The most frequent stationary vertices were the monomorphic nutrient specialists ($s_x = 0$ or $s_x = 1$). The second most-frequent stationary population composition was the dimorphism between two specialists ($s_x = 0$ and $s_x = 1$). Transitions

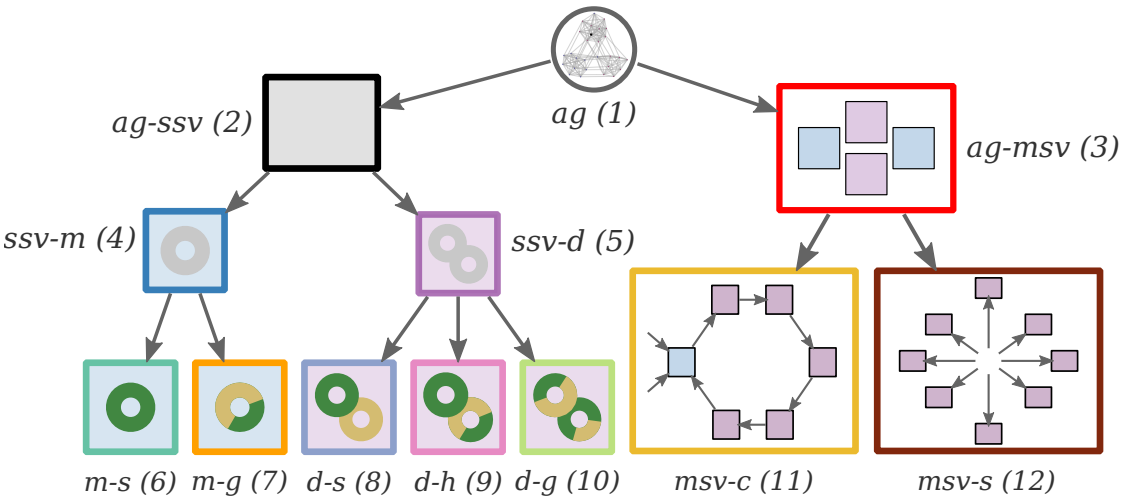


Figure 4.3.8: Labeling of graph stationary behavior. The labeling is hierarchical, although we will typically work with the leaves of the labeling tree (6-12). Abbreviations: *ag(1)*: all graphs, *ag-ssv(2)*: all graphs - single stationary vertex, *ag-msv(3)*: all graphs - multiple stationary vertices, *ssv-m(4)*: single stationary vertex - monomorphic, *ssv-d(5)*: single stationary vertex - dimorphic, *m-s(6)*: monomorphic - specialist, *m-g(7)*: monomorphic - generalist, *d-s(8)*: dimorphic - two specialists, *d-s(9)*: dimorphic - hybrid (a specialist and a generalist), *d-g(10)*: dimorphic - two generalists, *msv-c(11)*: multiple stationary vertices - cycle, *msv-s(12)*: multiple stationary vertices - stochastic.

between monomorphic vertices (i.e. invasion events where the mutant lineage outcompeted the parental lineage) were the most frequently observed across the ensemble, particularly at small maximum mutation norms.

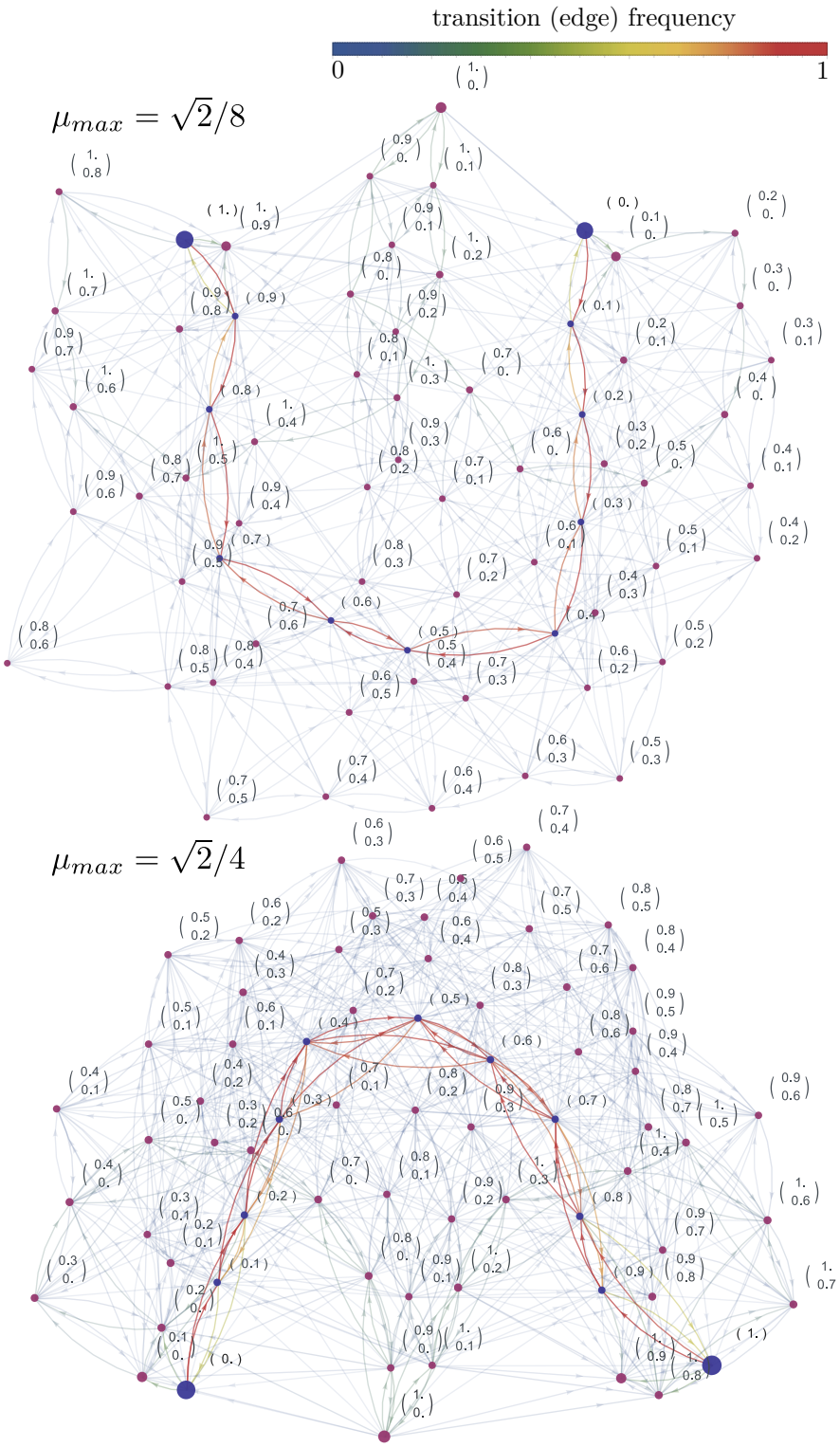
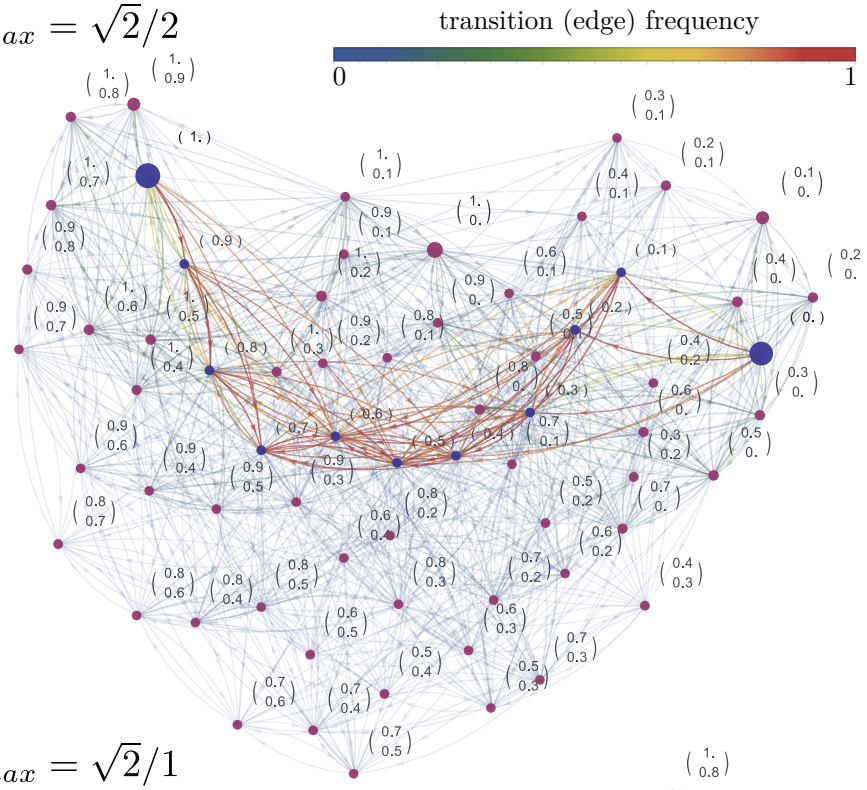
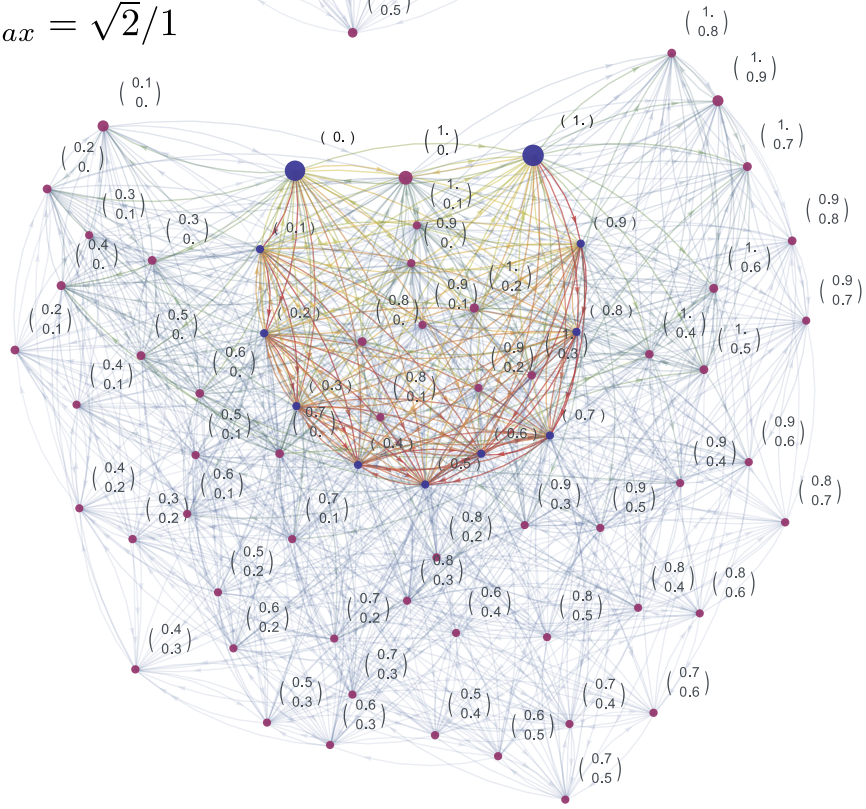


Figure 4.3.9: Average graphs from the 5000 random parameter sets. Vertices are labeled according to the resident phenotype(s). Vertex size is proportional to the mean per-vertex stationary probability across the ensemble. Blue vertices are monomorphic and purple vertices are dimorphic. An edge's color and opacity indicates the frequency with which it is observed across the ensemble - from light blue (no observations) to opaque red (observed in every graph). *Figure continues overleaf*

$\mu_{max} = \sqrt{2}/2$



$\mu_{max} = \sqrt{2}/1$



The nine model parameters can be reduced to five features by taking the log-transformed ratios of nutrients yields, nutrients concentrations, nutrient binding rates, and nutrient metabolic rates, together with the log-transformed dilution rate. Figure 4.3.10A shows the bivariate distribution of graph classes in a subset of the parameter space that was found to produce a reasonable separation of the classes.

To better describe the univariate association between model parameters and graph class (long-term stationary behavior) we used a K nearest neighbors density estimation method [**?**] to calculate the marginal probability of the graph class. Figure 4.3.10B shows the marginal probability of assigning a graph, with a particular value for each of the log-transformed model parameters, to each of the seven graph classes. We note, in particular, the strong dependence of large concentration ratios, and large dilution rates, on the dominance of the monomorphic specialist class. Graphs with multiple stationary vertices that exhibit a "stochastic" stationary probability distribution (high entropy) are found primarily in cases where the nutrient yields are equal. Graphs that exhibit limit cycle type behaviors are usually drawn from parameter sets where the log-transformed concentration ratio is approximately 3 – i.e. when there is a 1000-fold difference in the influx concentration of the two nutrients. The monomorphic generalist class is, in general, the least abundant of all the types of graphs and starts appearing when the yield difference is large (> 5).

The univariate association between model parameters and graph class provide some intuition about the parameters that are important in determining the stationary state. However, it has not been possible to find a combination of parameter transformations that produces a very good separation of the classes, even when we attempted to train classifiers with kernel-based methods. This suggests that the relationship between model parameters and graph stationary state is highly non-linear and very sensitive to small changes in parameter values.

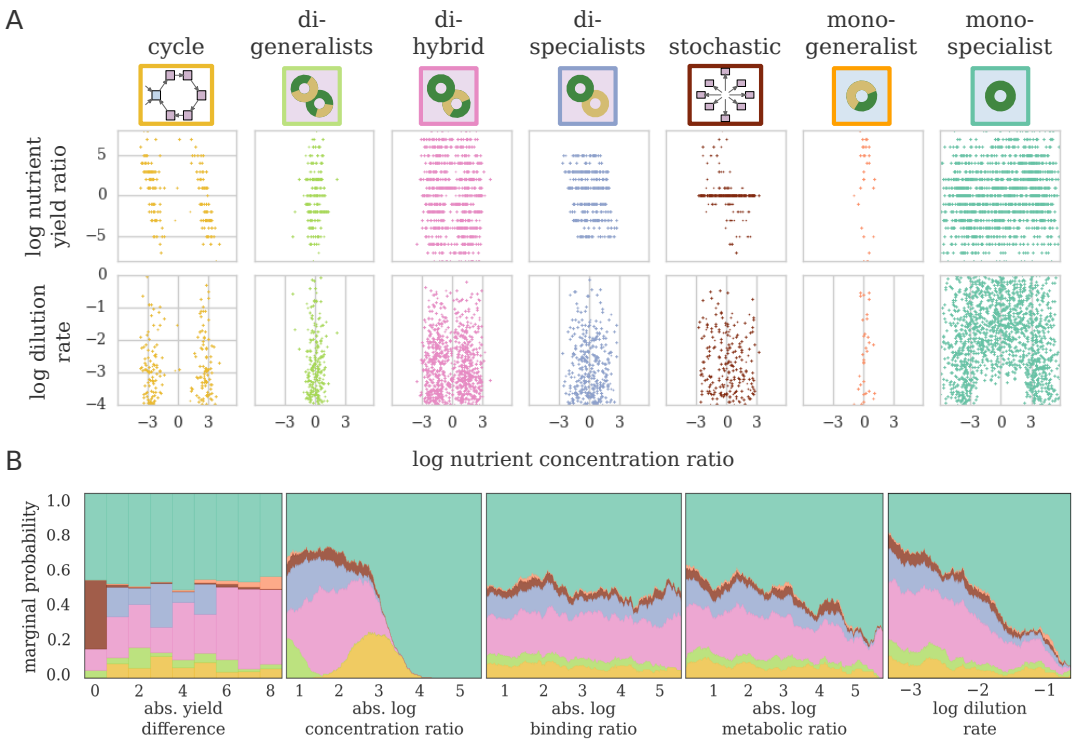


Figure 4.3.10: Association between model parameters and graph class (long-term stationary behavior). ***A***. Bivariate distribution of graph classes in a subset of the reduced model parameter space. The log-transformed nutrient concentration ratio separates the graph classes well. Note that the distributions are symmetric about the zero-axis of concentration and yield ratios. ***B***. A K nearest neighbors density estimation method, trained on univariate model parameter data, shows the marginal probability of assigning a graph to one of the 7 graph classes, for each value of a model parameter. The marginal probabilities are stacked such that the heigh of each color represents the corresponding class's probability.

4.3.2.1 Characterizing adaptation from graph measures

In the adaptation graphs produced by modeling approach, vertices correspond to dynamically-stable populations characterized by one or more phenotypes, and these are connected via edges if it is possible for a single mutant phenotype to transform the phenotype composition of the source vertex to that of the target vertex.

Because of the way we construct the adaptation maps we cannot calculate properties related to adaptation time (but see proposals in **??**). However, treating the graph as a discrete-time Markov chain, with "time" measured in terms of the number of mutations, we can nevertheless still calculate some interesting proper-

ties of adaptation (Figure 4.3.11).

For each graph in our data we enumerate all possible paths from the initial vertices (all monomorphic vertices) to the stationary vertex, or vertices. We calculate the probability of each such path as the product of the edge transition probabilities from the graph's adjacency matrix, weighted by the probability that the adaptation process begins in the first vertex of the path. For the cycle-type graphs, we stop the path when the process enters any vertex in the closed communication class.

$$P_{\text{path}} = P_{\text{init}}(v_0) \prod_{\text{edge}}^{\mathcal{E}} P_{\text{edge}} \tag{4.10}$$

We summarize the path length distribution for each graph by its mean and variance, and also record the number of unique paths. We can interpret the expected path length of a graph, with expectation taken over all unique paths, as the average number of mutation/invasion events. The variance in path length is interpreted as the variability we might observe across multiple parallel adaptation experiments. Finally, the number of unique paths is related to the repeatability of the adaptation path and the probability that two adaptation processes will realize the same phenotypic trajectory toward the stationary state.

We also determined the extent to which the adaptation process has bottlenecks on paths to the stationary vertices. To do this, we determined the size of the minimal vertex set that, when removed from the graph, disconnects an initial vertex from a stationary vertex. 'Cutting' a vertex from the graph means that we do not allow the adaptation process to assume the phenotype composition implied by the vertex; alternatively, we can remove all outgoing edges from the cut vertex, which is akin to protecting the phenotype composition from further adaptation. We applied this process for all combinations of initial and stationary vertices, and calculated the mean of the minimal vertex cut sizes, $|C|$, weighted by the probability that the process begins at the initial vertex, v_i, and ends at the stationary vertex v_j.

$$E[|C|] = \sum_i \sum_j P_{\text{init}}(v_i) P_{\text{stat}}(v_j) |C_{i,j}| \tag{4.11}$$

P_{init} is the initial probability of the process starting in vertex v_i, and P_{stat} is the

probability that the process will end in vertex v_j, assuming that it has started on v_i.

We interpret a low $E[|C|]$ as a bottlenecked graph. For example, a mean of 1 implies that the adaptation process can be prevented from reaching the stationary vertex by protecting a single vertex. On the other hand, a large mean cut means that the adaptation process has many alternative ways (vertex-independent paths) to reach the stationary vertex, and we would have to protect many vertices to arrest the adapting system.

Finally, we measured the sensitivity of the long-term stationary behavior to changes in model parameters. For each graph in our data we calculated the fraction of (30) nearest neighbors, in the reduced five-dimensional parameter space (each parameter standardized to zero mean and unit variance), that had a different stationary behavior compared to the focal graph. This dissimilarity fraction is a measure of the sensitivity of the stationary behavior to parameter changes, and is related to the controllability of the adaptation process. A large dissimilarity fraction implies that we can make a small parameter change to shift the current stationary behavior to a qualitatively different one – for example, if we have control over the chemostat parameters of the adapting systems. Indeed, we could train a classifier on our model parameter - graph class data, and request a new parameter set within some (standardized) radius of a focal point in parameter space that would shift the stationary behavior to a desired one. In small-scale tests, we obtained very good results in almost all cases of shifting from a current graph class to a target graph class (not shown).

We constructed the categorical distributions of the expected path length, the variance of the path length, the number of unique paths, the expected size of the minimal vertex cut, and the dissimilarity fraction, conditioned on each of the seven graph classes. We found that the maximum mutation norm had an important effect on the above graph measures, therefore we constructed the categorical distributions as a function of the maximum mutation norm (Figure 4.3.11). The distributions are, except in a few cases, relatively wide-spread, even after conditioning on the maximum mutation norm.

Using the median as a measure of a distribution's centrality, we found that graphs with a dimorphic specialist stationary vertex had the longest path length of all graph classes, at both small and large maximum mutation norms. This is per-

haps not very surprising since the two co-evolving phenotypes must move to both extremes of phenotype space. Therefore, we expect the largest number of mutations (between 10 and 15) for environments that support the co-existence of two specialist phenotypes.

We remark here on a few notable observations from Figure 4.3.11. For small mutation sizes, we found that graphs with a hybrid dimorphic stationary vertex had the highest median variance. Interestingly, we found that almost all graphs of this type had the *same* variance – in spite of a wide mean distribution – suggesting a conserved property of this type of graph. Variance in path length in general decreases sharply after a critical maximum mutation norm of approximately $e^{-1.32}$ (as do other graph measures that we calculated). Increasing the permitted size of mutations beyond this value allows a phenotype difference between mutant and resident of up to $||\Delta s|| = 0.33$, which introduces many new edges to the graph. This effect can also be seen in the number of unique paths, which jumps from approximately 20 before this critical value to between $e^5 \approx 150$ and $e^{10} \approx 22000$, depending on graph class. While the sharp increase is in part a side-effect of our phenotype discretization choice, we would still expect a smooth, but steep, increase in the neighborhood of this critical value for finer discretization choices. The dimorphic specialist class had the largest median number of unique paths (approximately 160000), while the monomorphic specialist class had the fewest. In general, all graph types have a single-vertex bottleneck when the maximum mutation norm is small. This is expected since mutations can only access only the immediate phenotypic neighbors of the residents. For the largest maximum mutation norms we found that the most bottlenecked type of graph is the stochastic one, followed by the limit cycle-type of graph (i.e. the multiple stationary vertices). The monomorphic generalist graph was the least bottlenecked, suggesting that it may be very difficult to prevent an adapting system from reaching such a steady state. Finally, we found that the monomorphic specialist stationary behavior is the most difficult to change qualitatively via a small parameter change, and therefore perhaps the most difficult to control in an experimental setting. In contrast, monomorphic generalist graphs were the most sensitive to model parameters and may be the most susceptible to environmental change.

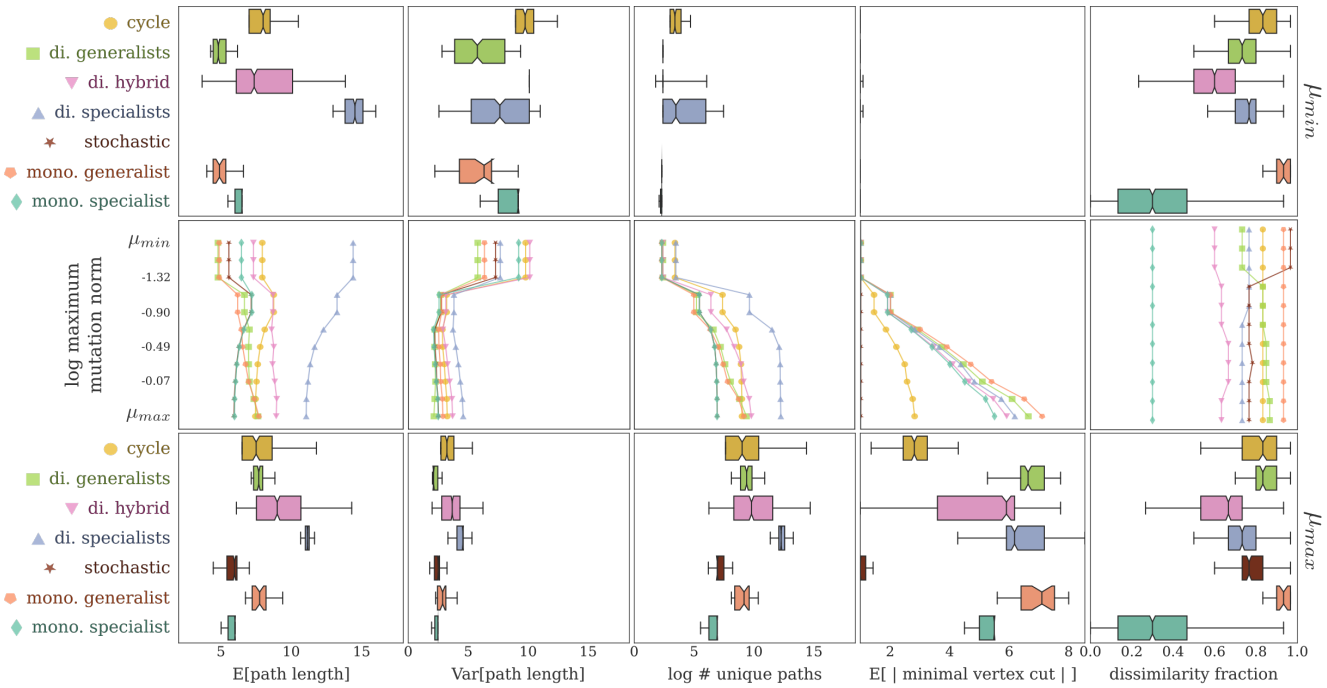


Figure 4.3.11: Summaries of aspects of adaptation calculated from graph measures. The categorical distributions, conditioned on the seven graph classes, are shown as boxplots for the smallest and largest mutation size (top and bottom rows). The median of the distributions, as a function of the maximum mutation norm, is shown as a line plot (middle row). We do not show the distribution for the stochastic-type graphs at the low maximum mutation norm extreme because there are very few such graphs in that regime. For each distribution, the box range shows the first and third quartile of the data, the interior band is the median and the notch size around the median denote the 95% confidence interval of the median. The whiskers denote the extent of the data range.

4.3.2.2 Graph centrality statistics predict long-term stationary behavior

In addition to the graph measures we discussed in Section 4.3.2.1, it is possible to summarize aspects of the adaptation graph topology by calculating various graph vertex centralities [138]. For each graph in our data we calculated standard centralities (in-degree centrality, out-degree centrality, closeness centrality, betweenness centrality, eigenvector centrality, and HITS (hub and authority centrality), and recorded the mean and variance of the distribution over the graph. We hypothesized that these summary statistics will be sufficient for detecting differences between graph classes, at least.

In general, we found that it was possible to train estimators on pairs of centrality statistic - graph class data to predict the graph class from centrality statistics alone. We used an ensemble classifier based on decision trees (random forest) for this problem. Briefly, tree-based classifiers iteratively partition the dataset by selecting the feature, and a point along that feature, that maximizes the reduction in class entropy after the selected point has been used to split the dataset. The class entropy of the parent is compared to the weighted sum of the post-split child node entropy to find the best possible feature split. Iterating this process results in a 'decision' tree that can be used to learn a classifier. Tree classifiers can be used in an ensemble whereby multiple 'weak' classifiers can be used together to generate a strong estimator from the majority vote of the constituent classifiers. Classifiers are typically weakened by only permitting the next split to be performed on a truncated subset of all possible features.

We used the ensemble classifier with 50 estimators and obtained a very good test set performance (> 0.95 macro-averaged, and cross-validated, F1 score). Given the strong discrimination capacity of the graph centrality statistics, we wanted to know whether it is possible to predict the long-term stationary behavior from incomplete explorations of the adaptation graph. An incomplete adaptation graph can be obtained, for example, from observations from multiple laboratory evolution experiments, each of which may or may not have reached the stationary vertex/vertices. We reasoned that, unlike the earlier graph measures (Section 4.3.2.1), that graph centralities may be more robust to incomplete graphs; consequently, they might provide more useful information about the long-term stationary behavior.

Toward this end, we randomly sampled each of the $55,000$ graphs in our data set to generate subgraphs containing a variable fraction of all edges (we called this fraction the graph 'coverage'). In total, we simulated 25 partial graphs at 9 graph coverage fractions (0.1 to 1.0), for each graph. After sampling the partial graph, we calculated the graph centrality statistics in the same way that we did for the complete graphs.

We then trained an estimator on complete-graph data (centrality statistic - graph label pairs) and evaluated the estimator's performance on partial graphs centrality data to assess if it is possible to predict the long-term stationary behavior of the graph – regardless of whether the stationary vertex was included in the partially

explored graph). We grouped the results according to the maximum mutation norm and graph coverage to highlight the regions where the estimator does best (Figure 4.3.12A). In general, classification performance increases as more of the adaptation graph becomes explored; however, we also found that partial-graph prediction is markedly better for larger mutation sizes. This observation is true for classifying complete graphs as well, though to a lesser extent. We hypothesize that graph centrality statistics converge more quickly to their complete-graph value for graphs with larger maximum mutation norms – i.e. they are more robust to missing data. We have not, however, tested this hypothesis yet.

It was possible to improve the performance of the classifier on partial-graph data by selecting subsets of the 14 graph centrality statistics to be used as estimator features. We trained and evaluated the performance of classifiers using all possible $2^{14} = 16384$ feature combinations. Figure 4.3.12B shows the distribution of classifier performance conditioned on graph coverage. The full-feature classifier (broken vertical lines) performs increasingly well, relative to the distribution of all classifiers, as the graph coverage approaches completeness – progressing from the left-tail of the classifier distribution to the right-tail at full coverage.

We reasoned that, particularly for sparse graph explorations, many graph centralities have not converged to their complete-graph value, and, as a consequence, might be providing misdirecting information to the estimator. Indeed, we find a correlation (not shown) between the number of centrality statistics used as features and the performance of the top classifiers at a specific graph coverage, with the ideal number of features increasing with increasing graph coverage.

In Figure 4.3.12C we show the performance of the classifier trained on the best-performing combination of features, optimized for each graph coverage value (column indicated by the black arrow at each heat map). Consistent with our earlier statement, the heat maps reveal the trade-off between prediction performance at low versus high graph coverage values. Figure Figure 4.3.12D shows the fraction of top 1% of classifiers, at each graph coverage value, that have included each of the 14 graph centrality statistics as a feature. It is difficult to provide an explanation with regards to why some centralities are preferentially used over others at specific graph coverage values – particularly for more complicated classification algorithms; however, we speculate that a centrality's utility as a feature depends both on its rate of convergence (as a function of graph coverage) as well as its

general discrimination utility.

The ideal strategy, then, for optimizing prediction is to use a subset of graph centralities based on the fraction of the adaptation graph that has been explored. It may be possible to infer the fraction of the adaptation graph that has been explored from a statistical analysis of the phenotype changes in the adaptation trajectories. The magnitude of the changes in phenotype space should decrease to zero as the graph approaches stationary state (the case of limit cycles is more complicated, but there the changes approach some small value). However, very sparse graphs will in general have many vertices with no out-going edges and the phenotype change will quickly decrease to zero. As the fraction of the graph that has been explored increases, trajectories will in general require more steps for the phenotype changes to decrease to zero. Figure 4.3.12E shows a characteristic result from analyzing the phenotypic change, at each mutational step, along trajectories constructed from a graph that has been explored to varying coverage. We find that, in general, the phenotypic change decays to zero more slowly as more of the graph becomes explored. This suggests that a statistical analysis (or machine-learning based approach) of partial graph trajectories might allow us to infer the graph coverage, and hence use the most well-suited classifier. We have not, however, pursued this analysis yet.

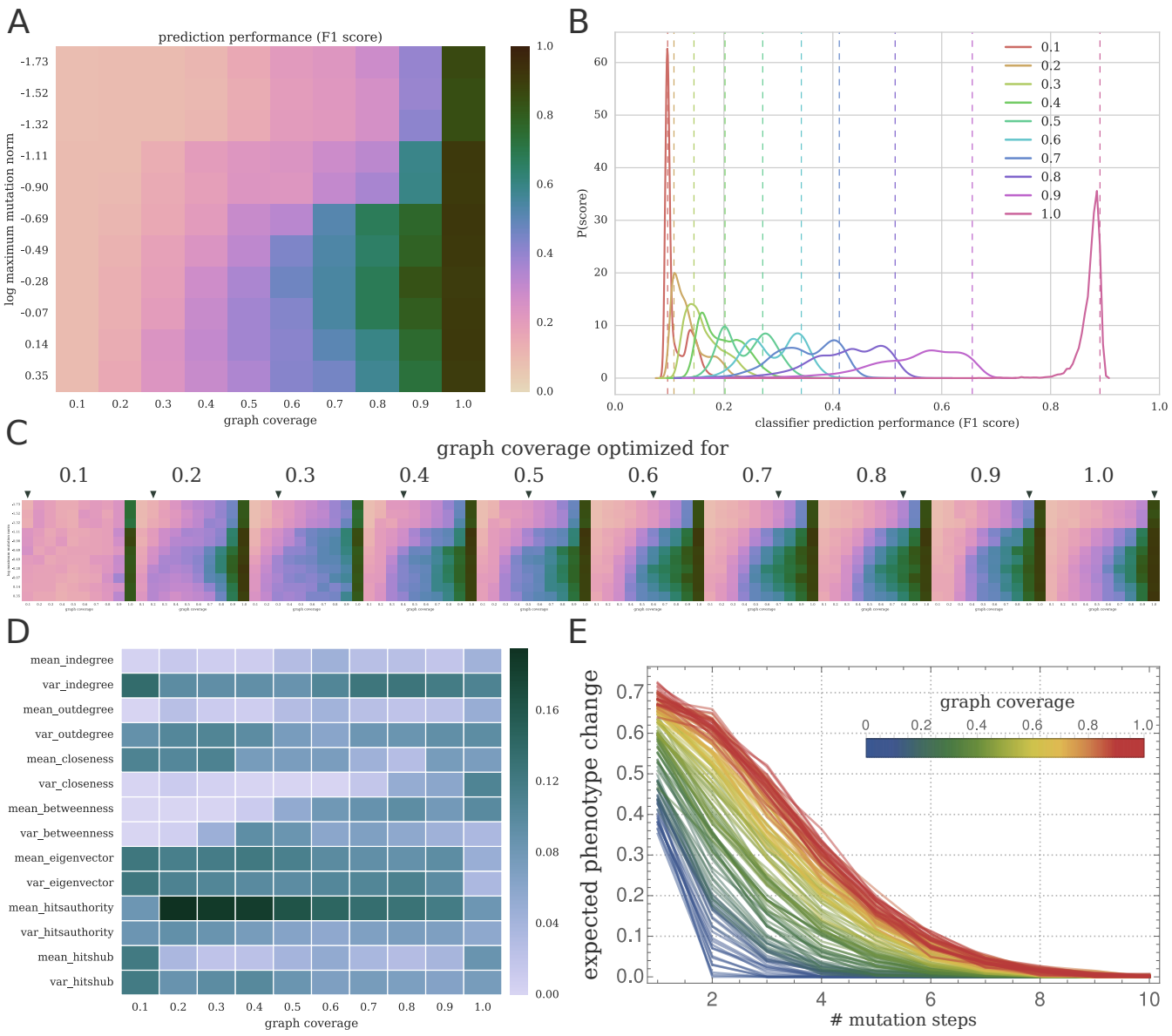


Figure 4.3.12: Predicting graph long-term stationary behavior from graph centrality statistics. **A** Classifier performance (F1 score, averaged over performance for each of the seven graph classes) grouped by maximum mutation norm and graph coverage (fraction of the graph explored). **B** Selecting classifier features reveals a trade-off between prediction at low coverage values versus prediction at high coverage values. Distributions show the classification performance of all 2^{14} centrality combinations, grouped by performance at a specific graph coverage value. The broken vertical lines indicate the performance of the full-set classifier. **C** Classification performance of the best classifier optimized for prediction at a specific graph coverage value. We note the trade-off between low-coverage versus high-coverage prediction performance. **D** Centrality inclusion as a classifier feature in the top 1% of classifiers optimized for performance at a specific graph coverage value. Values indicate the fraction of 1% classifiers that included the centrality statistic. **E** Inferring the fraction of the adaptation graph that has been explored, showing a characteristic result. The convergence of the expected phenotypic change, in terms of the number of mutational steps, is slower for graphs that have been more thoroughly explored. The expectation is taken at each mutational step, over many partial graphs (explored at a specific coverage value) and over many sampled trajectories on each graph. Colors indicate the graph coverage.

4.3.2.3 Graphs with multiple stationary vertices

We now turn our attention to the subset of adaptation maps that support graphs with multiple stationary vertices. Figure 4.3.13 shows the location of these graphs in a parameter subspace. Approximately 6% of the ensemble graphs, at the smallest maximum mutation norm, $\mu_{max} = \sqrt{2}/2^3$, had multiple stationary vertices, while approximately 12% of ensemble graphs at higher maximum mutation norms $(\sqrt{2}/2^2, \sqrt{2}/2, \sqrt{2})$ had multiple stationary vertices.

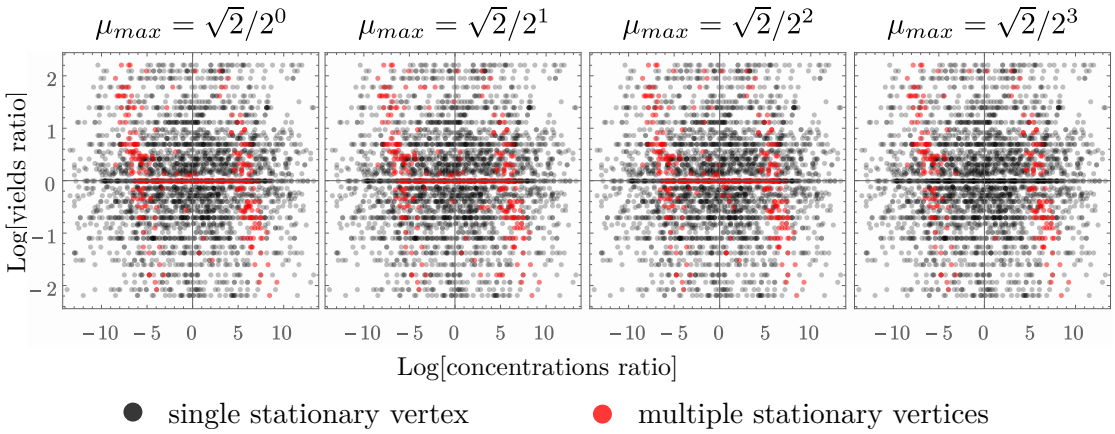


Figure 4.3.13: Location of maps with graphs that support multiple stationary vertices. The scatter plot of the log-transformed nutrient yields ratio versus the log-transformed nutrient concentrations ratio produces the most distinct organization and separation from the rest of the ensemble so it is shown here exclusively.

So far we have seen how multiple stationary vertices imply that the adaptation process can settle in one of many noninvasible population/phenotype compositions. However, another explanation for multiple stationary vertices is the presence of a set of communicating states within which the adaptation process cycles periodically. Figure 4.3.14 (bottom panel) shows the graphs from one representative parameter set. In this example, there are six stationary vertices that form a stationary communication set (Section **??**) that 'traps' the adaptation process; within this set, the system spends a some amount of time in each vertex. In other words, the population composition changes periodically between the six sets of phenotypes of the stationary vertices without achieving any *single* noninvasible phenotype composition.

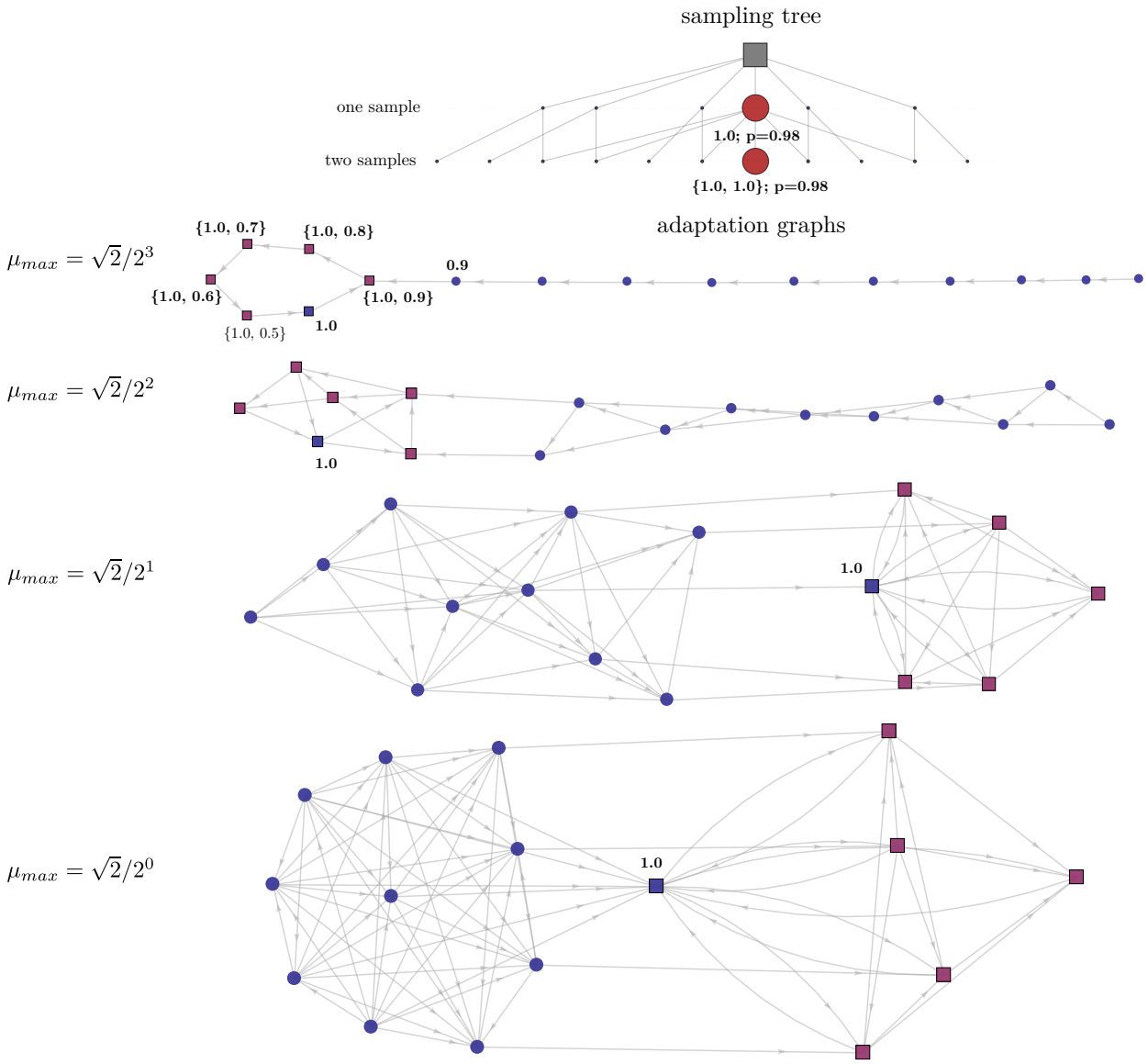


Figure 4.3.14: Example of parameter set with multiple stationary vertices. *Bottom:* Adaptation graphs for various maximum mutation norms, μ_{max}. In each case, there are six stationary vertices that form a closed, stationary communication class: five dimorphic vertices and one monomorphic vertex. The component phenotypes are indicated next to each vertex. All vertices have $s_x = 1.0$ in common. *Top:* Sampling probabilities for the phenotypes found in the stationary vertices. The first level of the tree shows the probability of sampling a phenotype when a single sample is taken, taking into account the vertex stationary probabilities as well as the fractional abundance that the phenotype establishes in each vertex/population composition. The second level shows the joint-probability of sampling two phenotypes when taking two independent samples (see text). The size of the nodes in the tree is proportional to the sampling probability. In this example, the sampling trees for each adaptation graph (bottom panel) were almost identical so only one is shown.

The adaptation process in Figure 4.3.14 takes a series of monomorphic steps (a sequence of single-phenotype population compositions) until a lineage with phenotype $s_x = 0.9$ emerges. This lineage can be invaded by $s_y = 1.0$ to establish a dimorphism (first stationary vertex in the communication 'loop'). Thereafter, a succession of phenotypes from $s_z = 0.8$ to $s_z = 0.5$ follows that each invade and displace the preceding non-specialist ($s \neq 1$) lineage. The last dimorphic vertex in the loop ($\{s_x = 1.0, s_y = 0.5\}$) is invaded by $s_z = 0.4$ and drives $s_y = 0.5$ to extinction; however, $s_z = 0.4$ is itself unable to establish a population at a high enough abundance (< 0.01 of the total population size) and is not retained during the mapping process, even though it out-competes $s_y = 0.5$ in the dimorphic resident environment in which it emerges. $s_x = 1$ remains as the sole resident until a lineage with phenotype $s_y = 0.9$ emerges again and the cycle re-iterates.

A closer examination of the resident lineages in the stationary vertices reveals that the specialist phenotype ($s_x = 1.0$) in fact dominates the population composition across all stationary vertices. Figure 4.3.14 (top panel) shows the probability of sampling a cell from each of the lineages in the stationary vertices as a way to summarize the phenotypes found in the stationary vertices while taking into account how much each phenotype is represented in the population composition of each vertex. This approach is particularly useful for interpreting graphs with multiple stationary vertices. The 'sampling tree' has two tiers after the root node: the first tier shows the probability of sampling a cell with a specific phenotype, while the second tier shows the joint probability of independently sampling two phenotypes. These probabilities are calculated as

$$\begin{aligned} P(S = x) &= \sum_{v \in \mathcal{V}_x} P_v \cdot f_v(S = x) \\ P(S = x, S = y) &= \sum_{v \in \{\mathcal{V}_x \cup \mathcal{V}_y\}} P_v \cdot f_v(S = x) \cdot f_v(S = y) \end{aligned} \tag{4.12}$$

for one and two independent samples respectively. P_v is the stationary vertex probability for vertex v calculated according to the discrete Markov process analysis (Section **??**) and some initial distribution for the ancestral phenotype - for example, the uniform distribution over the monomorphic vertices. $f_v(S = x)$ is the probability of sampling a cell with phenotype x given the population composition of stationary vertex v, and is just the fraction of the population that has

phenotype x. The sum is taken over all stationary vertices that contain phenotype x, in the single-sample case, and over the union of stationary vertices that contain x or y in the two-sample case.

One interpretation of the above calculation is that it answers the following question: Suppose that the adaptation process was started randomly, according to our initial assumptions about the ancestral phenotype, and given enough time to reach stationarity. What phenotypes would we observe if we were to randomly sample one or more cells from the chemostat vessel? Alternatively, if multiple parallel adaptation experiments were run to stationarity, and we sampled each chemostat - taking one or more cells - how frequently would we observe a particular phenotype or combination of phenotypes?

In the above example the specialist phenotype ($s_x = 1.0$) fixes in the chemostat environment and establishes a mutational landscape where small and transient successions of variant phenotypes appear in the population; however, the likelihood of observing these small populations is very low.

Not all graphs with multiple stationary vertices are represented by the preceding example. Figure 4.3.15 shows an example from the second class of graphs that were found in the parameter sampling ensemble. These graphs typically, though not always, have distinct stationary vertices that do not belong in a single communication set. Multiple stationary vertices, in these cases, imply that the adaptation process has a higher degree of stochasticity and historical contingency - where the set of mutations, and the sequence of mutations, ultimately direct the process to one of many stationary states. These graphs are not dominated by a single phenotype, though one or two phenotypes may be more frequently observed across multiple adaptation experiments.

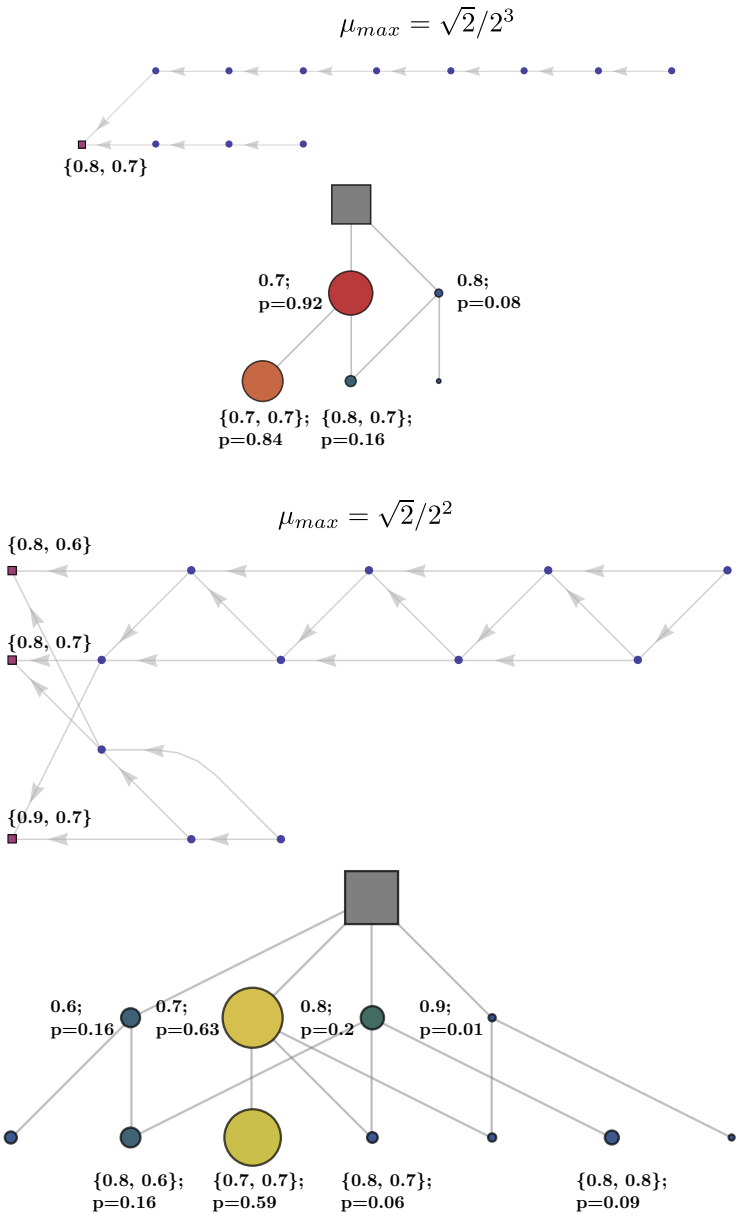
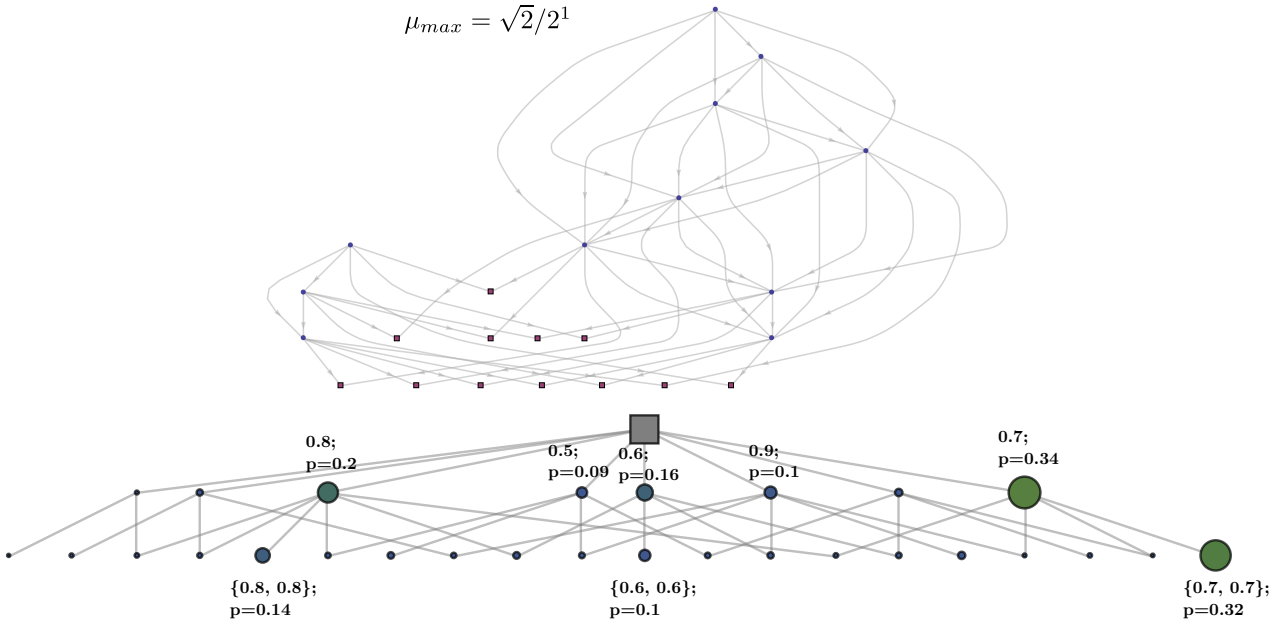
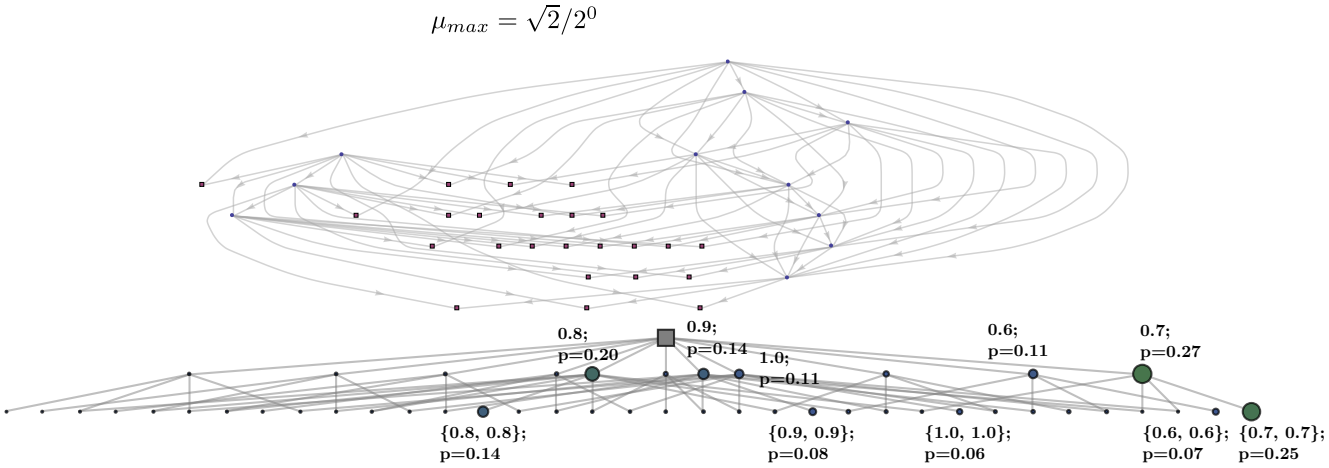


Figure 4.3.15: Representative example of the second class of graphs with multiple stationary vertices, showing the adaptation graphs, together with their associated sampling trees, for four values of the maximum mutation norm, μ_{max}. In each adaptation graph square vertices are stationary and are colored blue if they are monomorphic and purple if dimorphic. Each sampling tree has two levels: the first level shows the probability of sampling a particular phenotype that is found in (any one of) the stationary vertices, while the second level shows the joint probability of independently sampling two phenotypes (equation 4.12). Each node in the tree represents the outcome of one or two sampling events. The initial distribution used for process analysis was uniform on the monomorphic vertices. The highest sampling probabilities in each tree level are indicated next to their node. *Figure continues overleaf...*


$\mu_{max} = \sqrt{2}/2^1$
0.8;
p=0.2
0.5;
p=0.09
0.6;
p=0.16
0.9;
p=0.1
0.7;
p=0.34
{0.8, 0.8};
p=0.14
{0.6, 0.6};
p=0.1
{0.7, 0.7};
p=0.32


$\mu_{max} = \sqrt{2}/2^0$
0.8;
p=0.20
0.9;
p=0.14
1.0;
p=0.11
0.6;
p=0.11
0.7;
p=0.27
{0.8, 0.8};
p=0.14
{0.9, 0.9};
p=0.08
{1.0, 1.0};
p=0.06
{0.6, 0.6};
p=0.07
{0.7, 0.7};
p=0.25

In this particular example the number of stationary vertices increases with the value of the maximum mutation norm, from a single vertex ($\mu_{max} = \sqrt{2}/2^3$) to 24 vertices ($\mu_{max} = \sqrt{2}$) for the largest value where every phenotype can appear in a mutant lineage regardless of the phenotype of its parent. The stationary phenotypes are peaked around $s_x = 0.7$, with $\{0.7, 0.7\}$ having the highest sampling joint probability for all maximum mutation norms; however, the distribution becomes less peaked with increasing μ_{max} (last graph/tree in Figure 4.3.15). Interestingly for this example, even though all of the stationary vertices are dimorphic, the probability of sampling two cells with different phenotypes is small relative to the probability of observing two cells with the same phenotype. This outcome arises from the combination of the stationary vertex probabilities and the fractional compositions of the population in each vertex, which are not necessarily dominated by a single phenotype.

The two examples shown here represent the two types of adaptation maps whose graphs contain multiple stationary vertices. The first type is characterized almost always by a single communication class of stationary vertices over which the process is periodic. This class is dominated by a single phenotype; as a consequence, the highest joint probability very large (close to 1). The second type of graph does not always have a single communication class. Rather, it may consist of a combination of closed 'loops' and distinct stationary vertices and, as a result, the highest joint probability is usually not close to 1.

Figure 4.3.16 shows the distribution of the two types of multiple stationary vertex (MSV) graphs in a parameter subspace. There are two points to note. First, most MSV type-two (historical contingency) graphs are occur when the two nutrients yields are equal; in contrast, type-one (cyclical) graphs arise in two narrow bands of roughly-linear relationship between the (absolute) log yields ratio and (absolute) log concentration ratio. Second, MSV type-two graphs (historical contingency) are only possible above a threshold maximum mutation norm ($\sqrt{2}/2^3$), whereas type-one graphs emerge for all maximum mutation sizes. We can interpret the distribution of MSV type-two (historical contingency), and to a lesser extent of the MSV type-one (cyclical) graphs, as the conditions that promote increased variability in the long-term stationary behavior of adapting systems.

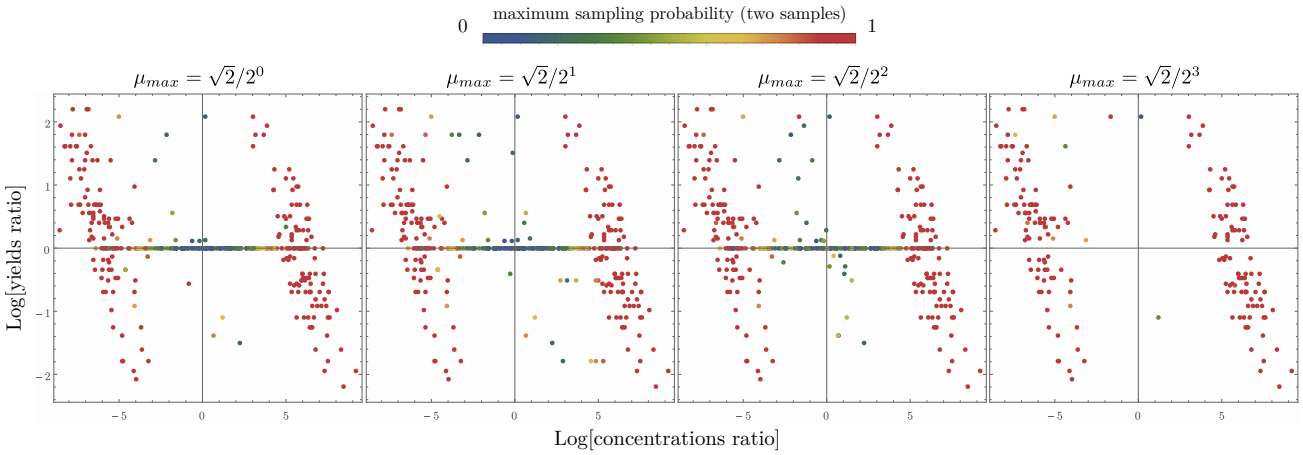


Figure 4.3.16: Location of the two types of adaptation maps with multiple stationary vertices in the log-transformed nutrient yields ratio versus the log-transformed nutrient concentrations ratio subspace. The maximum joint-sampling probability is close to 1 (red points) for maps of the first type (single stationary communication class), whereas for the second type of map it is usually between 0 and 1 (blue-yellow points).

4.3.3 Section summary and discussion

In this section we have defined a resource-consumer model with two (perfectly) substitutable nutrients. A structured population was used to implement a phenotype-determined nutrient 'preference', whereby a cell's capacity to bind and metabolize one of two environmental nutrients was parametrized as an adaptable trait (Equation 4.5).

We demonstrated the utility of numerical methods for this analytically intractable model. In particular, we showed how we can generate a comprehensive map of adaptation dynamics for models where there is no simple expression for the invasion fitness, and when the outcome of competition cannot easily be known from prior measurements (Section 4.2.1).

We then conducted an exploration of the model's parameter space to understand how parameters might affect the adaptation process, with an emphasis on stationary outcomes. We have found that approximately 70% of sampled parameter sets had graphs with a single stationary vertex. The remaining 30% had graphs that supported multiple stationary vertices.

Approximately half of the graphs with a single stationary vertex had a monomorphic specialist stationary outcome ($s_x = 0$ or $s_x = 1$). Very few (approximately

1%) of these graphs had stationary generalist outcomes (Table **??**). Interestingly, however, generalist phenotypes were the most common response to a resident specialist phenotype (61% of dimorphic vertices contained a specialist and a generalist). Only approximately 22% of dimorphic vertices, on the other hand, consisted of two specialists ($s_x = 0, s_y = 1$). We conclude that it is possible, under certain conditions, that the modification of the environmental conditions by a resident lineage/phenotype allows phenotypes that were otherwise not adaptively stable strategies to establish and co-exist in the population. Our results are in agreement with the theoretical predictions of Ballyk and Wolkowicz [136], who also found cases where one lineage cannot survive in the absence of a competitor.

Parameter sets that produced graphs which supported multiple stationary vertices could be divided into two classes. On the one hand, we found a class of graphs whereby the adaptation process had a periodic stationary behavior involving multiple communicating vertices (Figure 4.3.14). In these cases, a dominant lineage could be invaded by a phenotype to establish a weak coexistence (mutant fixes at low relative abundance). The rare mutant could then be invaded by another mutant, and so on..., establishing a succession pattern. At the end of the cycle, a mutant lineage invades and drives the previous rare co-resident lineage to extinction - but does not itself establish at a significant concentration. This means that a monomorphic population remains, after the resolution of the invasion event, consisting of only the dominant phenotype. Then the sequence of invasions begins again. Such cyclical patterns of invasions are the hallmark of game theory and frequency-dependent fitness effects [65, 66, 67], and are highly reminiscent of the 'rock-paper-scissors' models from economic game theory [139].

The second class of multiple-stationary vertex graphs (Figure 4.3.15) consisted of disconnected stationary vertices, unlike the previous class. In these cases, the outcome of the adaptation process is truly probabilistic. The initial (ancestral) phenotype, together with the precise history of mutations, ultimately determines the stationary phenotypes. These results are in agreement with theoretical and experimental evidence regarding the effect of historical contingency and chance in adaptive evolution [140, 141]. In general, such effects manifest as both decreased repeatability in adaptation trajectories as well as increased variance in the final outcomes of adaptation across replicate laboratory evolution experiments [142, 143].

4.4 Three substitutable nutrients

In this section we will present some results from a model of adaptation on three substitutable nutrients. We have followed the same methodology as in the model with two substitutable nutrients, since the two models are similar.

The model with three nutrients is an incremental step in complexity compared to the one in Section 4.3. In particular, the reaction channels (shown schematically in Figure 4.3.1 and in Equation 4.5) have been preserved; in addition, one new set of nutrient influx, nutrient uptake, and metabolism reactions have been included for the new nutrient type (Table 4.3 shows the collected model parameters).

symbol	parameter
y_1	nutrient 1 yield
k_1	nutrient 1 binding rate
m_1	nutrient 1 metabolic rate
p_1	nutrient 1 concentration
y_2	nutrient 2 yield
k_2	nutrient 2 binding rate
m_2	nutrient 2 metabolic rate
p_2	nutrient 2 concentration
y_3	nutrient 3 yield
k_3	nutrient 3 binding rate
m_3	nutrient 3 metabolic rate
p_3	nutrient 3 concentration
D	dilution rate

Table 4.3: The 13 parameters for the model with three substitutable parameters. Each nutrient is associated with a yield, binding rate, metabolic rate, and concentration parameter. The dilution rate is the same for all nutrients (related to the constant volumetric flow out of the chemostat vessel), and is the same for the model's cell populations.

The adaptation parameter lies on the unit simplex in $\mathbb{R}^3$ (Figure 4.4.1) and parametrizes a cell's 'preference' for one species of nutrient relative to the others; more mechanistically, it can be interpreted as the fraction of cell-surface nutrient

transporters specialized for internalizing a specific type of nutrient (or group of nutrients).

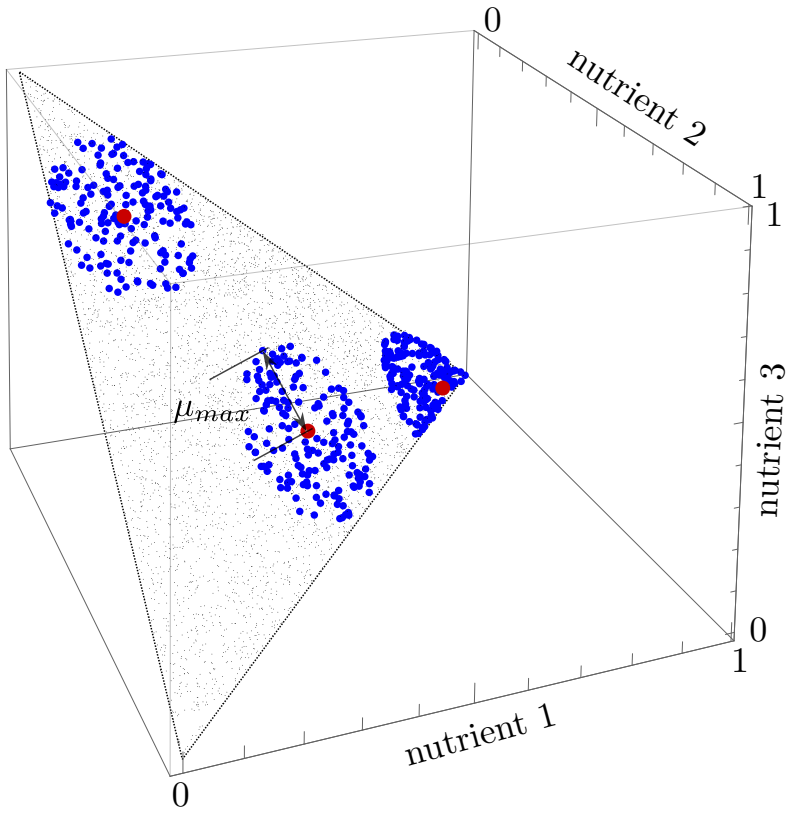


Figure 4.4.1: The unit simplex encloses the adaptation parameter space. The maximum mutation norm, μ_{max}, is the maximum allowed euclidean deviation of a mutant phenotype (blue points) from its progenitor phenotype (red points) - shown here for three example progenitor phenotypes. The adaptation/mutation kernel describes the distribution of mutant phenotypes around the progenitor (we have only used the uniform distribution in this chapter).

4.4.1 Triple co-existence in an example parameter set

As with the previous section, we will begin with a demonstration parameter set (Table 4.4) to illustrate the added complexity that a third nutrient brings to the problem. As anticipated, with three substitutable nutrients the maximum co-residency number is now increased to three. Because the number of possible resident-mutant invasion models increases as the power of this maximum co-residency number, the unit line was discretized into fewer elements (five, for a total of 625 possible vertices) to keep the mapping process computationally inexpensive.

symbol	parameter	value
y_1	nutrient 1 yield	8
k_1	nutrient 1 binding rate	5.9
m_1	nutrient 1 metabolic rate	0.049
p_1	nutrient 1 concentration	0.39
y_2	nutrient 2 yield	7
k_2	nutrient 2 binding rate	0.0024
m_2	nutrient 2 metabolic rate	260
p_2	nutrient 2 concentration	2.6
y_3	nutrient 3 yield	5
k_3	nutrient 3 binding rate	0.85
m_3	nutrient 3 metabolic rate	0.42
p_3	nutrient 3 concentration	0.12
D	dilution rate	0.0074

Table 4.4: Parameter values for the three-nutrient example.

The adaptation graphs, stationary vertex distributions, and sampling trees are shown in Figure 4.4.2. When mutations are unbounded ($\mu_{max} = \sqrt{2}$) the only stationary vertex contains three phenotypes: $\{0.5, 0.5, 0\}, \{0.0, 1.0, 0.0\}, \{0.0, 0.0, 1.0\}$; that is, there are three lineages, the first of which uses the first and second nutrient species equally, the second lineage is a 'specialist' of the second nutrient, and the third lineage is a 'specialist' of the third nutrient. The second phenotype/lineage is the most abundant ($P = 0.39$ of sampling three cells with the $\{0.0, 1.0, 0.0\}$ phenotype), followed by the first phenotype. The third lineage establishes at a small fraction of the total population composition and is relatively rare. For the smallest maximum mutation norm, $\mu_{max} = \sqrt{2}/4$, the graph has four stationary vertices, with $\{0.5, 0.5, 0.0\}, \{0.25, 0.25, 0.25\}, \{0.0, 1.0, 0.0\}$ having the highest probability. The $\{0.0, 1.0, 0.0\}$ remains the most abundant one, with $P = 0.39$ of sampling three cells of this types across all stationary vertices. Finally, in contrast to the larger μ_{max}, the $\{0.25, 0.25, 0.50\}$ phenotype can establish at an intermediate fractional abundance, in at least one stationary vertex, and there is a 0.11 probability that a sample of three cells from the stationary process will contain one cell with this phenotype.

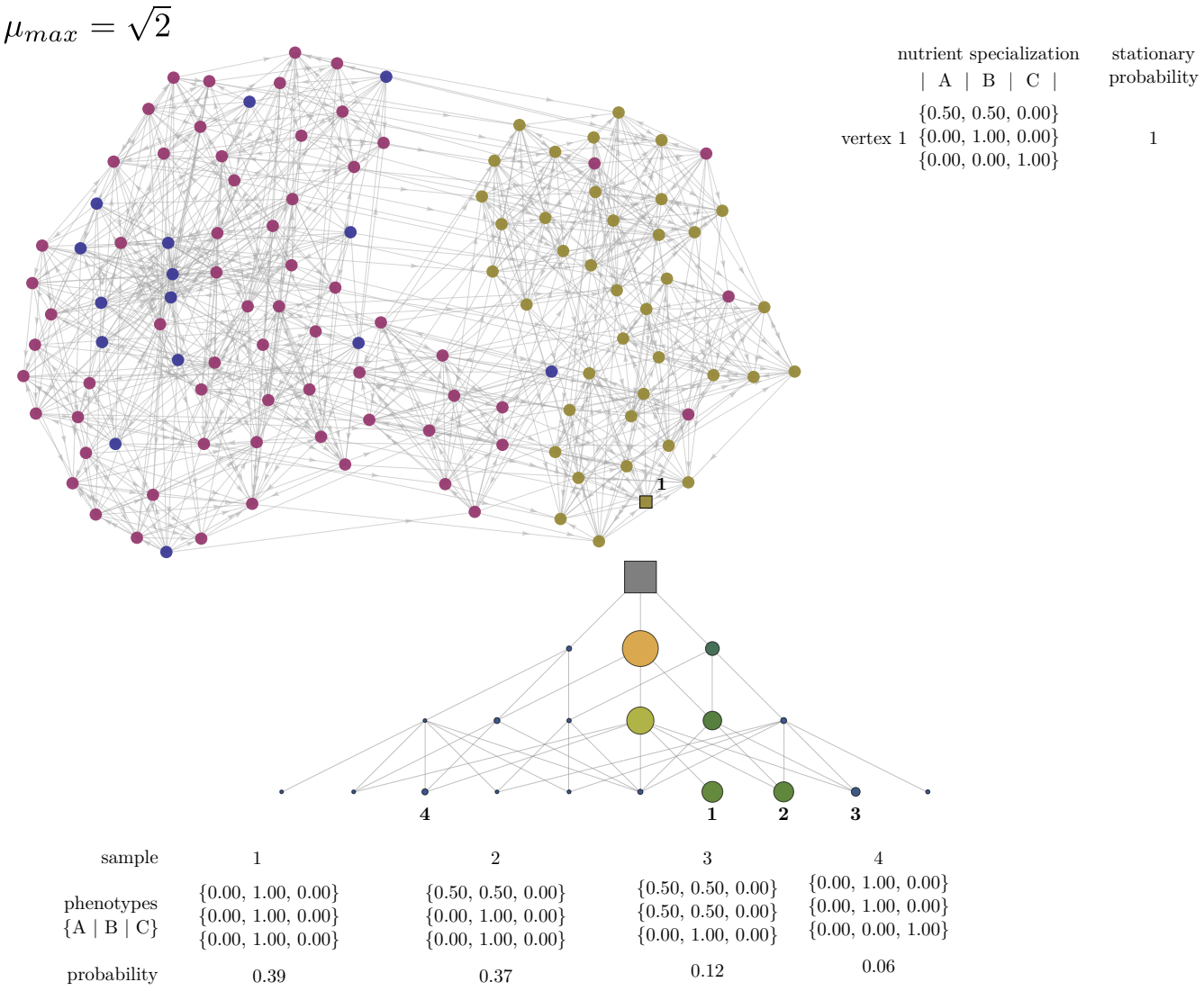
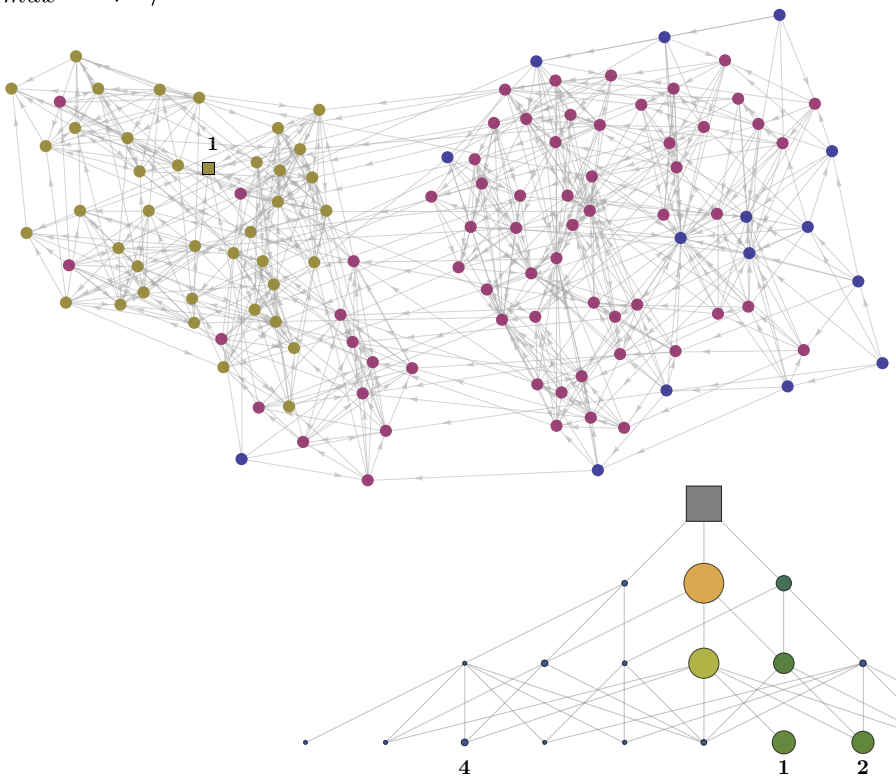
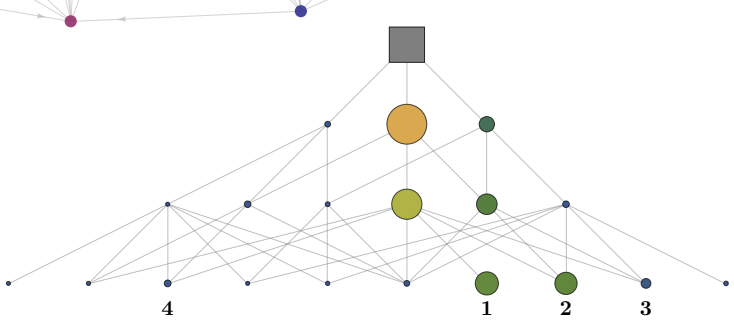


Figure 4.4.2: Adaptation graphs and sampling trees for the example parameter set using three maximum mutation norm values. In the adaptation graph, blue vertices are monomorphic, purple vertices are dimorphic, and yellow vertices have three resident phenotypes. The stationary vertices and stationary vertex probabilities are indicated to the right of each graph and the five highest joint probabilities of sampling three phenotypes, according to the sampling tree, are shown at the bottom of each panel. The uniform initial distribution on the monomorphic vertices was used for stationarity calculations. *Figure continues overleaf...*

$\mu_{max} = \sqrt{2}/2$

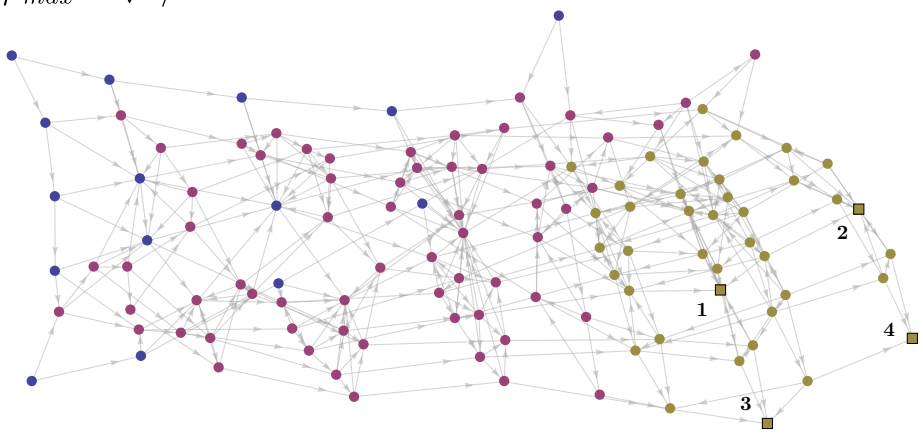


	nutrient specialization | A | B | C |	stationary probability
vertex 1	{0.50, 0.50, 0.00} {0.00, 1.00, 0.00} {0.00, 0.00, 1.00}	1

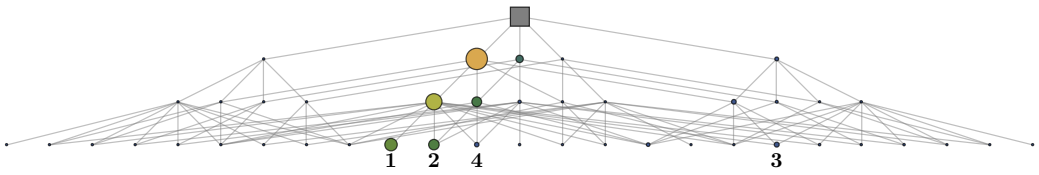


sample	1	2	3	4
phenotypes {A | B | C}	{0.00, 1.00, 0.00} {0.00, 1.00, 0.00} {0.00, 1.00, 0.00}	{0.50, 0.50, 0.00} {0.00, 1.00, 0.00} {0.00, 1.00, 0.00}	{0.50, 0.50, 0.00} {0.50, 0.50, 0.00} {0.00, 1.00, 0.00}	{0.00, 1.00, 0.00} {0.00, 1.00, 0.00} {0.00, 0.00, 1.00}
probability	0.39	0.37	0.12	0.06

$\mu_{max} = \sqrt{2}/4$



	nutrient specialization | A | B | C |	stationary probability
vertex 1	{0.50, 0.50, 0.00} {0.25, 0.25, 0.25} {0.00, 1.00, 0.00}	0.925
vertex 2	{1.00, 0.00, 0.00} {0.25, 0.25, 0.50} {0.00, 1.00, 0.00}	0.007
vertex 3	{0.50, 0.50, 0.00} {0.00, 1.00, 0.00} {0.00, 0.00, 1.00}	0.0674
vertex 4	{1.00, 0.00, 0.00} {0.00, 1.00, 0.00} {0.00, 0.00, 1.00}	<0.001



sample	1	2	3	4
phenotypes {A | B | C}	{0.00, 1.00, 0.00} {0.00, 1.00, 0.00} {0.00, 1.00, 0.00}	{0.50, 0.50, 0.00} {0.00, 1.00, 0.00} {0.00, 1.00, 0.00}	{0.25, 0.25, 0.50} {0.00, 1.00, 0.00} {0.00, 1.00, 0.00}	{0.50, 0.50, 0.00} {0.50, 0.50, 0.00} {0.00, 1.00, 0.00}
probability	0.39	0.32	0.11	0.088

4.4.2 Parameter space sampling and global analysis

Continuing with the methodology followed in previous sections, we randomly sampled the parameter space to generate $1,000$ adaptation maps and construct their graphs. The average graphs for three μ_{max} values were computed and are shown in Figure 4.4.3. Table 4.5 summarizes the connectivity of the average graphs in terms of the vertices' co-residency number (monomorphic, dimorphic, trimorphic).

Number of edges between labeled vertices ($\mu_{max} = \sqrt{2}/4$)

	monomorphic	dimorphic	trimorphic	total (outgoing)
monomorphic	11271	3911	0	15182
dimorphic	8116	26355	3155	37626
trimorphic	0	1027	12535	13562
total (incoming)	19387	31293	15690	

Number of edges divided by number of (target) vertices ($\mu_{max} = \sqrt{2}/4$)

	monomorphic	dimorphic	trimorphic	total (outgoing)
monomorphic	751.4	37.2	0	788.6
dimorphic	541.1	251	8.1	800.2
trimorphic	0	9.8	32.1	41.9
total (incoming)	1292.5	298	40.2	

Table 4.5: Graph connectivity between vertices labeled according to their co-residency number for the average graphs of the ensemble. The $\{i, j\}$th entry is from co-residency level i to co-residency level j. *Table continues overleaf...*

Number of edges between labeled vertices ($\mu_{max} = \sqrt{2}/2$)

	monomorphic	dimorphic	trimorphic	total (outgoing)
monomorphic	22413	10905	0	33318
dimorphic	14745	43474	6960	65179
trimorphic	3	1515	17971	19489
total (incoming)	37161	55894	24931	

Number of edges divided by number of (target) vertices ($\mu_{max} = \sqrt{2}/2$)

	monomorphic	dimorphic	trimorphic	total (outgoing)
monomorphic	1494.2	103.9	0	1598.1
dimorphic	983	414	17.8	1414.9
trimorphic	0.2	14.4	46.1	60.7
total (incoming)	2477.4	532.3	63.9	

Number of edges between labeled vertices ($\mu_{max} = \sqrt{2}$)

	monomorphic	dimorphic	trimorphic	total (outgoing)
monomorphic	30127	21770	0	51897
dimorphic	16799	49398	10023	76220
trimorphic	3	1608	18858	20469
total (incoming)	46929	72776	28881	

Mean number of edges per labeled (target) vertex ($\mu_{max} = \sqrt{2}$)

	monomorphic	dimorphic	trimorphic	total (outgoing)
monomorphic	2008.5	207.3	0	2215.8
dimorphic	1112	470.5	26	1616.1
trimorphic	0.2	15.3	48.4	63.9
total (incoming)	3129	693.1	74.1	

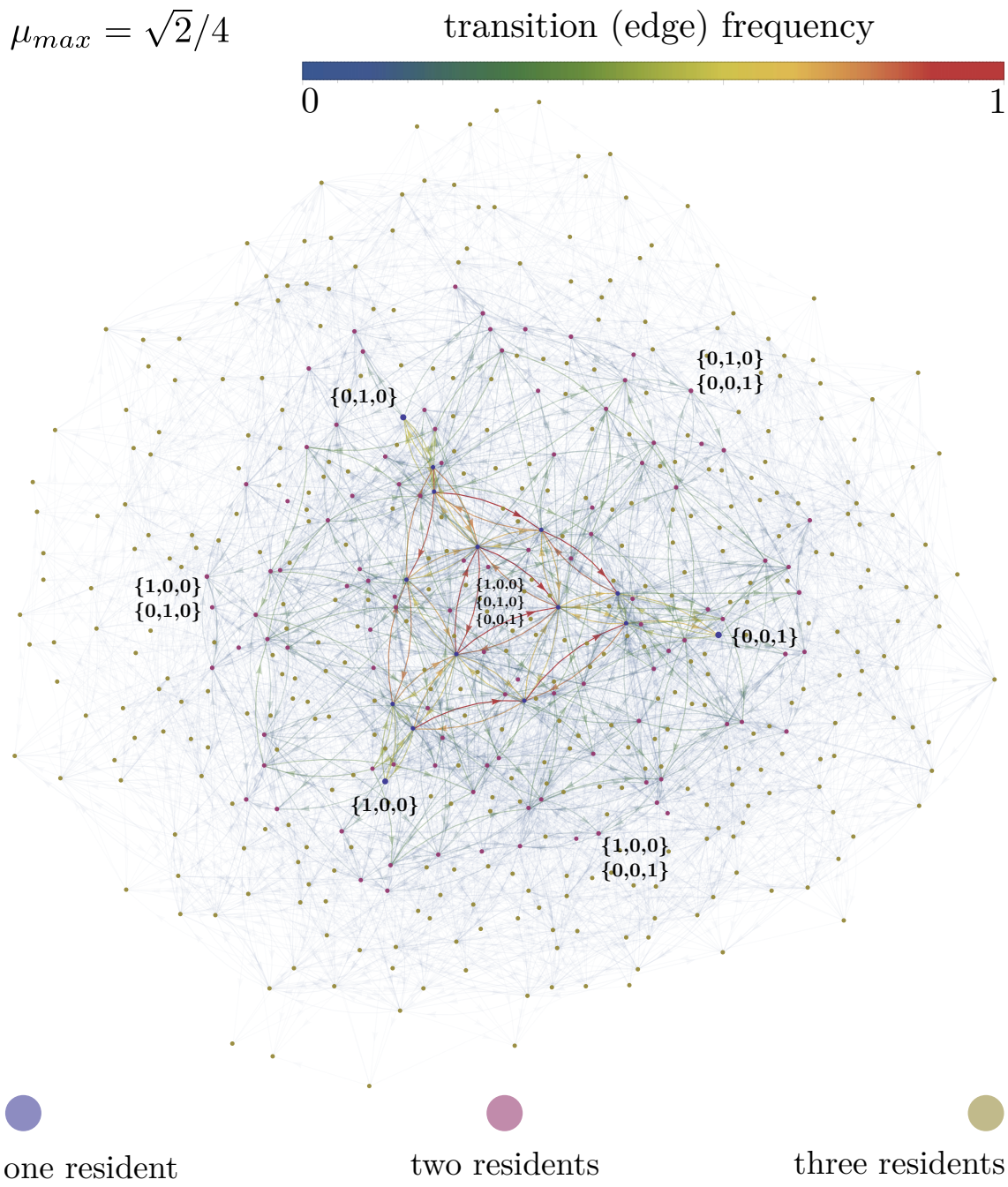
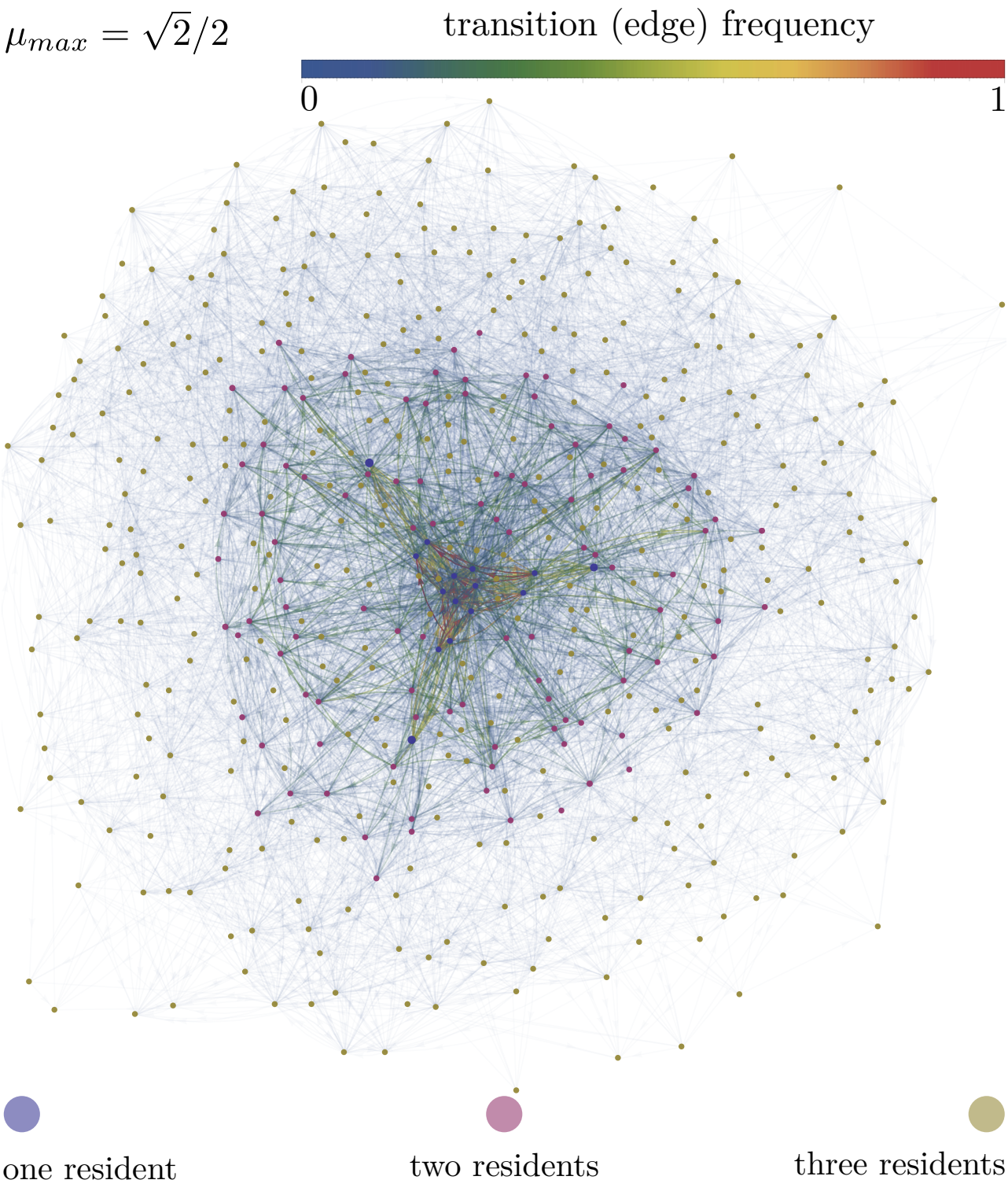
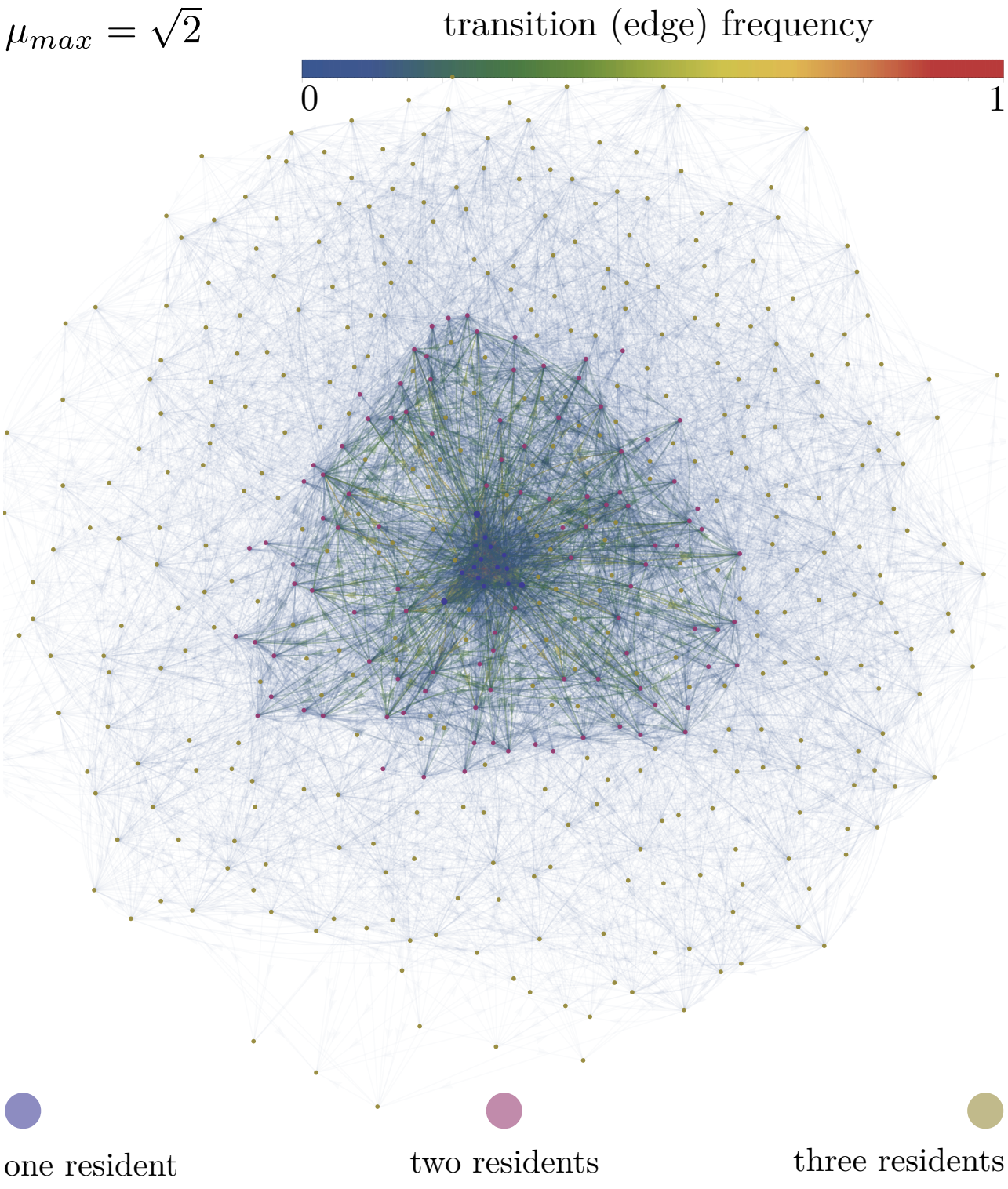


Figure 4.4.3: Average graphs for the ensemble of 1,000 adaptation maps. Blue vertices are monomorphic, purple vertices are dimorphic, and yellow vertices are trimorphic. The size of each vertex is proportional to the mean per-vertex stationary probability across the ensemble. The color and opacity of each edge is proportional to the frequency with which it appeared in the ensemble. *Figure continues on the next two pages...*


$\mu_{max} = \sqrt{2}/2$
transition (edge) frequency
0
1
one resident
two residents
three residents


$\mu_{max} = \sqrt{2}$
transition (edge) frequency
0
1
one resident
two residents
three residents

The most frequently observed type of transition across the ensemble, for all maximum mutation norms, was between a dimorphic vertex and another dimorphic vertex - i.e. a resident population consisting of two lineages is invaded by a nascent mutant lineage that establishes and drives one of the two resident lineages to extinction. This was then followed by transitions from monomorphic vertices to monomorphic vertices, and from trimorphic vertices to trimorphic vertices. If the number of transitions into the k-morphic vertices is divided by the number of vertices, as a crude estimate of the expected number of incoming edges per vertex (the in-degree of the vertex), then we obtain a clear ranking of monomorphic $>$ dimorphic $>$ trimorphic, which is in agreement with the observation that the per-vertex stationary probabilities also rank in a similar manner.

In general, diversification events, whereby a nascent mutant establishes but does not drive the resident to extinction, become more common as the maximum mutation norm increases (since descendants can reach phenotypic niches that are farther away from the parent); in addition, once diversification occurs, it was relatively uncommon, though not rare, for it to be lost - both from one to two lineages and from two to three lineages - as a result of invasion by a novel phenotype.

In Appendix A the average graphs are re-interpreted by highlighting (resizing) the graph vertices proportionally to three centrality metrics (vertex degree, vertex betweenness, and vertex closeness) that can be used as measures of the 'importance' of a vertex in the graph. As mentioned already, the monomorphic vertices have the highest vertex degree (number of incoming and outgoing edges); that is, except for the specialist vertices ($\{1,0,0\},\{0,1,0\},\{0,0,1\}$) which have few outgoing connections. The dimorphic vertices are the most important for 'information' flow through the graph - as measured by the betweenness centrality metric, which is based on the number of shortest paths through a vertex. While dimorphic population compositions may not be as likely to be adaptively stable, they nevertheless act as important intermediates in the adaptation process. The trimorphic vertices had the highest closeness centrality values, from which we can infer that they are important for the quick spread of 'information' through the graph. As with the dimorphic vertices, they are even less likely to be adaptively stable, and graphs are not usually disconnected when trimorphic vertices are removed; however, the adaptation process transitions through populations with three co-resident lineages in shorter paths to the stationary phenotype com-

positions.

Understanding the effect of model parameters on the stationary behavior of adaptation graphs is a more challenging task when there are three substitutable nutrients, not least because,the parameter space can't be reduced to a feature space that can be easily visualized. We now need two pairwise comparisons between matching nutrient-specific parameters (yield, concentration, binding rate, and metabolic rates) to describe the relative parameter values. Other approaches to obtaining a good feature space, such as the mean deviation of the parameter set's grouped parameters from each other, or the mean deviation of the parameter set's grouped parameters from the ensemble's mean for that group, did not yield good results.

Figure 4.4.4 shows one feature space that gave the best separation between graphs with monomorphic, dimorphic, and trimorphic stationary vertices. Graphs with a single stationary vertex that contain three phenotypes have nutrient concentrations that are more similar to each other - i.e. no one nutrient is much more abundant than the others - compared to graphs with monomorphic and dimorphic stationary vertices. Dimorphic graphs and monomorphic graphs have nutrient concentrations that may be more dissimilar, with the dimorphic class having a smaller covariance, indicating that the trend for polymorphism is 'inward' toward zero in this feature space. Moreover, trimorphic graphs require the lowest dilution rates, followed by dimorphic graphs, which are followed by the monomorphic graphs which appear exclusively for the highest dilution rates. The chemostat dilution rate controls the population growth rate since, at steady state, the rate of population loss must equal the rate of population growth; therefore, in fast-growing environments - for example in environmental conditions (temperature, pH, oxygen concentration, etc...) that are conducive to organism growth - it is usually easier for a single phenotype to dominate the environment, whereas harsher environments wherein organisms grow slowly are more conducive to stable phenotypic diversity. Figure 4.4.5 shows the classification of a random forest classifier trained on the ensemble data using the log-transformed ratios of matched nutrient parameters and the log-transformed dilution rate (13 features total). The classifier returns higher classification probabilities for trimorphic graphs at the lower end of the dilution rate range and for small nutrient concentration ratios, high probabilities for dimorphic graphs at intermediate dilution rates and medium nutrient concentration ratios, and high probabilities for monomorphic graphs at

low dilution rates and large nutrient concentration ratios.

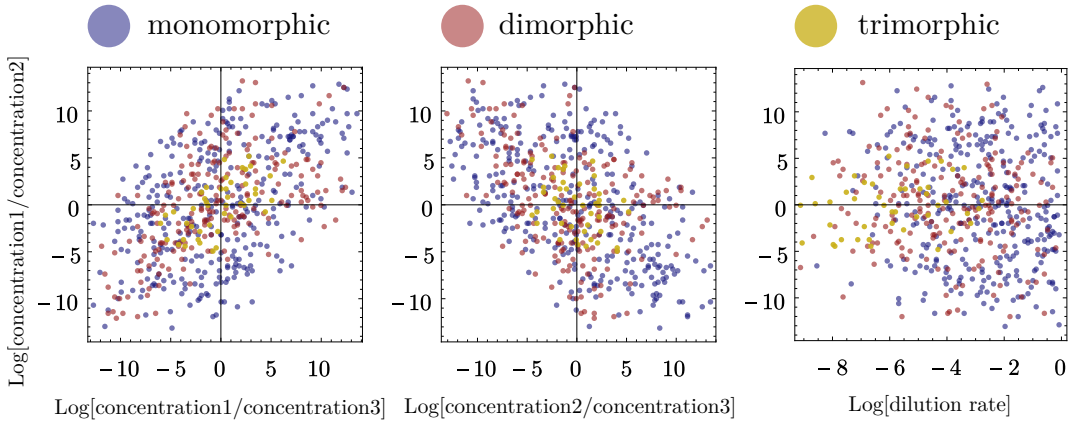


Figure 4.4.4: Parameter space transformation to a reduced-dimensionality feature space. The log-transformed nutrient concentration ratios versus each other, and versus the log-transformed dilution rate, gave the best class separation.

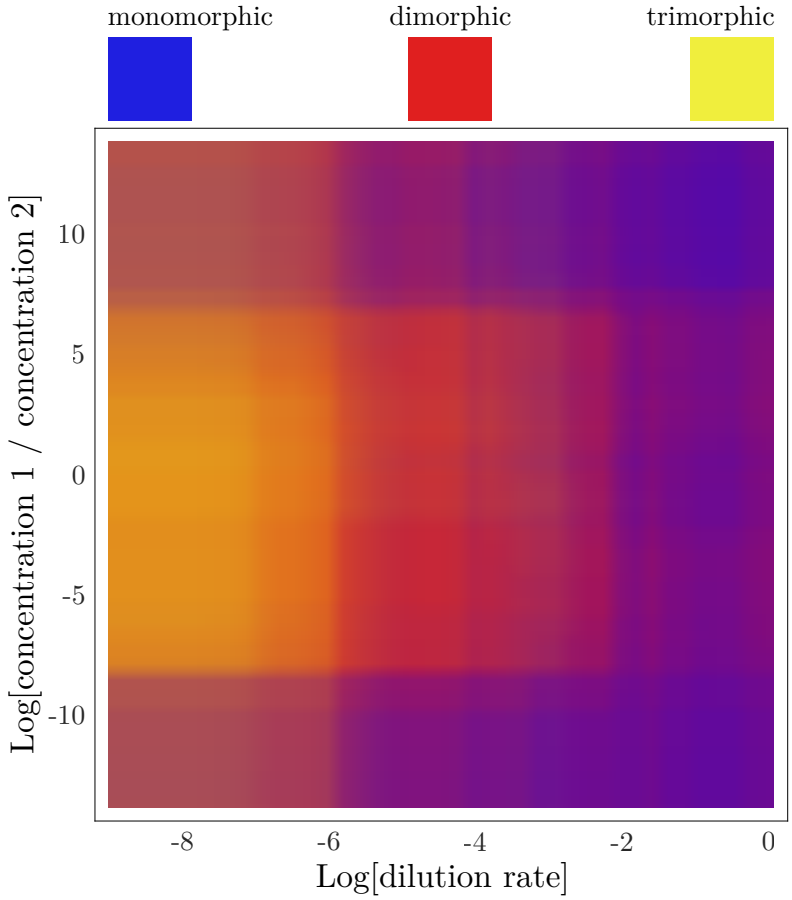


Figure 4.4.5: Classification probabilities from a random forest classifier where the blue (monomorphic), red (dimorphic), and trimorphic (yellow) colors have been blended in proportion to the classification probability.

4.4.3 Section summary and discussion

In this section we have investigated a model of adaptive dynamics where cells grow on three, perfectly substitutable, limiting nutrients.

Using an example parameter set we anticipated the stability of triple co-existence (Figure 4.4.2), which has been made possible because of the inclusion of the new nutrient. We again demonstrated the importance of the maximum mutation norm for the determination of the stationary vertex probabilities.

We then examined the adaptation processes for an ensemble of model parameter sets. We computed the adaptation maps for randomly sampled parameters and constructed the resulting graphs using several maximum mutation norms. We inspected and categorized the outcomes of invasion events and showed how diversification events, whereby the number of co-resident lineages/phenotypes in the population increases, become more common as the maximum mutation norm increases. This is not unexpected [53,56,144]: a mutant lineage with nutrient-use requirements that are sufficiently dissimilar from those of the parental lineage will face decreased competition (from the parent) and is more likely to survive. We also observed that events that decrease diversity are relatively uncommon, though not exceedingly rare (Table 4.5). Finally, we found that, in chemostat-like environments, low dilution rates were associated with higher diversity, whereas high dilution rates - where cells must grow fast to survive being washed out of the chemostat environment - usually restricted the number of adaptively stable co-resident lineages (Figures 4.4.4 and 4.4.5).

4.5 Applications summary

In this chapter we have applied the numerical methods described in Chapter **??** to three models of adaptive dynamics.

In Section 4.2 we revisited the model with a single limiting nutrient and two alternative metabolic pathways. We demonstrated how the numerical approach extends the standard analysis of adaptive dynamics to provide a more complete description of the adaptation process. Going beyond the invasion condition and invasion fitness, we showed how the parameters of the mutation kernel, together with the initial (ancestral) phenotype distribution play an important role in determining the stationary outcome of adaptation.

In Section 4.3 we conducted a numerical investigation of a model with two perfectly substitutable nutrients. The analytical intractability of the model provided a good candidate to fully demonstrate the scope and utility of the numerical approach. In particular, using adaptation maps, graphs, and discrete Markov process analysis, we presented a comprehensive investigation of the possible adaptive dynamics that the model can support throughout its parameter space. We hierarchically classified the adaptation graphs according to their stationary behavior and determined where each class localizes in a feature subspace produced by simple transformations of the model parameter space.

In Section 4.4 we extended the model with multiple nutrients to three perfectly substitutable nutrients. Following a similar approach, we demonstrated that a dynamically stable co-existence between three lineages/phenotypes is possible, and that it may be the adaptively stable phenotypic composition. Sampling the parameter space and surveying the resulting adaptation graphs, we presented results on the type and frequency of phenotype composition changes and found a parameter space transformation that enables a classifier to learn the co-residency level (diversity) in the stationary state of the adaptation process.

Chapter 5

Conclusions and outlook

Adaptive evolution is the process by which natural selection, acting on phenotypic variation within a population, promotes the survival of phenotypes that are more successful at contributing to future generations. Adaptation takes place at the intersection of population genetics, natural selection, and underlying constraints that restrict the range of possible phenotypes [11, 12, 13]. Moreover, the adaptation process involves feedbacks from multiple and disparate timescales, such as demographic, competition/invasion, and mutation/adaptation timescales.

Modeling adaptive evolution has a history that begins with the mathematical analysis of human and animal populations [36, 37]. The formulation of adaptive dynamics theory standardized and formalized an approach focused on describing and predicting the trajectories of phenotypic characteristics through time [68, 69, 70, 71]. Unlike previous efforts, it provided a tool set within which ecological realism did not have to be sacrificed for the sake of theoretical rigor.

Nevertheless, adaptive dynamics theory contains its own limitations, for which it has been criticised [76]. Notable among these are the commonplace assumptions that 'mutant' phenotypes are not very different from those of their immediate progenitors and that the phenotype space is continuous. Where these assumptions have been relaxed, analyses of the resulting adaptation processes have not deviated substantially from the type of analysis developed for the standard theory; as a consequence, an incomplete description of the modes and scope of adaptive evolution has emerged and much of the methodology remains fragmented.

In this work we have presented a set of modeling concepts and numerical methods

that, building on the theory of adaptive dynamics, provide one way in which the process of adaptation can be modelled and understood more comprehensively.

Our work was motivated by questions regarding the effect of mechanistic constraints on adaptive evolution [85, 86]. In particular, we wished to investigate the role of universal cellular trade-offs - such as the partitioning of ribosomal machinery [97], the cell surface area, and generally the allocation of finite cellular resources between competing processes [96, 99] - in shaping the outcomes of the adaptation process.

In this chapter we will summarize what we have done, synthesize and discuss new information, and present avenues for future research.

5.1 A framework for modeling adaptive dynamics

In Chapter **??** we developed a modeling framework that integrates ecology, constrained phenotypes, and adaptive dynamics. Progressing incrementally from standard chemostat-like models, we made extensive use of population structure to implement decision points that allow cellular resource allocation strategies to determine how cells grow and interact with the environment. We derived invasion conditions for the various models that are in agreement with previous theoretical work [40, 42, 45]. The modeling framework we described is generally applicable to situations where phenotypes are constrained. By parametrizing the phenotype on a (unit) simplex, we allow a lineage's strategy to determine the probability with which mutually exclusive cellular actions are taken. Moreover, because these decisions are taken at the ecological level, it is possible to empirically measure and characterize a species/lineage's phenotype from laboratory and field observations [96]. Our modeling approach should be readily applicable when phenotypes are found to involve mutually exclusive decisions, and can generate predictions about the future phenotypic trajectory. The models could also potentially extrapolate backwards to reconstruct possible phenotypic histories, provided that the environmental conditions have not changed substantially - although, in Chapter **??** we present findings that suggest how chance and stochasticity can complicate such a task.

In Section ?? we found that the (invasion) fitness of a metabolic pathway preference strategy/phenotype can be decomposed, under some conditions, in terms of the fitness of individual cell growth trajectories. Because heterogeneity among clonal populations of cells is widespread [145, 146], it may be difficult to quantify the overall (invasion) fitness of the population from measurements prior to actual competition. In this work we have shown how we can combine the fitness of individual cells, which can be tracked over long time periods using microfluidic systems and quantitative microscopy [147], to derive an expression for a lineage's invasion fitness. In addition, we have demonstrated one way in which heterogeneity may arise, stemming from the ecological decisions of individual cells, which in turn leads to the short-term fitness differences among an otherwise clonal population. Selection may act on these fitness differences to promote metabolic regulation through which cells can further optimize real-time pathway choice, according to their growth state and/or past metabolic decisions. Such regulation is the norm in living cells and our theoretical results provide evidence that it would indeed emerge from adaptive evolution starting with simpler, more naive, metabolic strategies.

5.2 A tool set for the analysis of adaptive dynamics models

In Chapter ?? we described a set of numerical methods for generating, simulating, visualizing, and analyzing models of adaptive dynamics. Our approach centers on discretizing the phenotype space and simulating all possible dynamical models that might arise from competitions between phenotypes. We presented a dynamic-programming algorithm to construct adaptation maps that significantly decreased our simulation times. Because we do not require an analytical invasion condition - relying, instead, on a large volume of automated simulations - our numerical methods are well-suited to handle models that are analytically intractable and for which a closed-form expression for the invasion condition may not exist [40]. Because of the discretization of the phenotype space, we have also addressed a long-standing criticism of adaptive dynamics theory that calls into question the continuity of possible phenotypes. We highlight how discretization can indeed introduce barriers in the adaptation process, that may ultimately

change stationary outcomes, by eliminating stable intermediate phenotype compositions in an otherwise smooth process.

We then demonstrated the utility of our numerical methods by applying our methodology to three models of adaptive dynamics in Chapter **??**. Starting with the model with two alternative metabolic pathways, we first compared the results from the numerical approach with those from Chapter **??**, which are more like the standard approaches using adaptive dynamics theory [73]. We demonstrated how the invasion condition, simple as it may be, cannot, by itself, comprehensively represent the adaptation process - although it can be used to generate such representations. In particular, we found that the manner in which new (mutant) phenotypes are derived from a parental progenitor can change the dynamics and stationary outcomes of adaptation; sometimes, this change is quite dramatic, as with our example illustrations where both metabolic (exclusive) specialists are adaptively stable strategies but the stationary probabilities of each strategy depend on the size of the maximum mutation norm. The repeatability of adaptive evolution is a topic of much historical and contemporary interest [19, 32] and our results now provide evidence that the degree of variability in dynamical and stationary outcomes may depend on the mutational statistics - in other words, on the availability and supply of novel phenotypes. In addition, we have demonstrated the importance of the initial (ancestral) phenotype distribution on determining stationary outcomes. Historical contingency is another important topic in microbial evolution, whose role in adaptive evolution is actively being researched [29]. Our results provide evidence in support of an influential role for such randomness - although we cannot yet generalize beyond a proof of concept, since our results show that the degree to which historical contingency affects the adaptation process must be treated on the basis of individual cases.

Our investigation of the adaptation of nutrient preference in environments with multiple perfectly substitutable nutrients fully exploited the capabilities of the numerical methods. Without a universal expression for the invasion condition (which has eluded other authors [137]) we are not able to analytically predict the outcome of competition during an invasion event. Our methods provide one possible way in which to organize and systematically conduct the appropriate simulations in order to derive a complete picture of the adaptation process.

We conducted a random sampling of the models' parameter space to investigate

how these parameters affect adaptation, particularly with regard to the stationary outcome. We then used a hierarchical classification approach to group the graphs of the adaptation processes according to their stationary outcomes. We found that some phenotypes are better off in the presence of a competitor phenotype - in the sense that the combination of the two phenotypes is adaptively stable whereas any one phenotype on its own would be prone to invasion. This finding lends support to theoretical predictions for such a possibility from other work [136]. To be more precise, we find that nutrient generalists - i.e cells that do not use a single nutrient exclusively - are more likely to be maintained in a population that already contains a nutrient specialist. This, and other related results, are the hallmark of frequency-dependent selection [33, 67], whereby the fitness of a phenotype is conditioned on the types of other phenotypes already present in the population. This result provides a framework with which the experimental observations of parallel evolutionary dynamics of microbial metabolic diversification can be understood [35].

We also found that adaptation graphs can, under certain conditions, support multiple stationary vertices - i.e. multiple adaptively stable phenotype compositions. Notably, we found that in some cases the graph's stationary vertices fall into a single communication class (group of vertices). The adaptation process becomes trapped in this class and cycles periodically between dynamically stable, but invasible, phenotype compositions, which are usually dominated by a single phenotype. We find that this 'rock-paper-scissors' type of cyclical successions [139,148] is relatively common, occurring in a localized region of the model's parameter space, and supports experimental observations of such behavior [149]. Finally, we find that stable phenotype diversity is greater in low dilution environments, where the cell population grows slower at steady state. Conversely, environments with a high dilution rate, where cells must grow faster to survive, do not usually support multiple co-residents. Because the natural analog of a chemostat's dilution rate is mortality - from both intrinsic (i.e ageing) and extrinsic (i.e predation) sources - we can conclude that the variability in adaptively stable strategies will in general be lower in high-dilution environments compared to those environments where the cells do not have to grow fast in order to survive dilution and washing-out.

5.3 Validating models of adaptation with microbial adaptation experiments

In this section we will discuss how our results, in particular the results from the two-nutrient model in Chapter 4 (Section **??**), can be used to make predictions that can be tested experimentally. We will outline one possible experimental setup that can generate test data, describing the required genetic constructs, data generation, and how results will be compared with theory. The approach we describe does not rely on laboratory evolution; instead, we generate genetic constructs that parametrize an organism's preference between two substitutable nutrients and carry out competitions between the strains.

Our goal is to experimentally approximate the evolutionary 'tuning' of an organism's nutrient preference between two substitutable nutrients. Evolution in natural systems will act on many agents involved in the utilization of nutrients in the cell to tune this preference. To simplify our study of adaptation, however, we can focus on a single point in the nutrient utilization process that, other things being equal, can be understood as setting the organism's nutrient preference. The experimental procedure will involve constructing strains with a parametrizable relative nutrient preference between two substitutable nutrients, and testing all possible competition/invasion scenarios in continuous-growth environments for a variety of environmental parameters. Competition results will then be used to construct the experimental adaptation graph that can be directly compared to the one generated from our model. The competition experiments can be conducted in parallel using multiplexed chemostat arrays [150] for replication and reducing waiting time between experiments, since the environmental (chemostat) parameters are constant for all competitions.

5.3.1 Genetic constructs and nutrient parameters

The genetic constructs necessary for this experiment will modify the nutrient transporters. The objective is to make strains with different constitutive levels of transporters for each nutrient. One possible way to achieve this is to replace the endogenous promoter of a transporter with inducible promoters (such as the GAL4 promoter) or by utilizing an auxin-based degron system for targeted trans-

porter degradation [151]. Since we will use chemostat-like environments, we can set the transcription level by flowing in a constant concentration of the inducer. Unfortunately, this method requires the use of many orthogonal components (one for each competing strain per nutrient). Alternatively, we can 'discretize' the transporter concentration range by replacing the endogenous transporter promoters with promoters of varying strength [152]. In our parameter exploration section for the two-nutrient model (Section **??**), for example, we used 10 discrete phenotypes so the number of genetic constructs required to test the theory is not too large. The constructs should also be genetically barcoded [153] to enable identification and quantification after competition experiments to resolve the invasion outcome. The sequencing reactions can be multiplexed, with appropriate tags, so the results can be gathered in batches.

In accordance with our model, competitions between phenotypic strains will take place in a chemostat-like environment. In a standard chemostat setup, the experimenter has control over parameters that appear in our model, such as the dilution rate, volumetric flow rate, and the concentrations of the incoming nutrients. The effect of environmental parameters on adaptation can be investigated experimentally, in a manner similar to our parameter exploration approach (Section **??**).

The remaining parameters in our model are concerned with nutrient metabolism, namely the uptake rate, metabolic rate, and yield. The yield of a nutrient (ϕ and θ in our model) can be calculated from the mass of organism produced from a certain mass of nutrient:

$$\text{nutrient yield} = \frac{\text{mass of organism}}{\text{mass of nutrient}} \tag{5.1}$$

at the end of a batch-growth experiment, after all of the nutrient has been consumed. The yields of the two substitutable nutrients in our experiment can then be converted to the discrete yield terms required by our model. For example, if $\gamma_u = 0.2$ and $\gamma_v = 0.4$ are the nutrient yields then we can set $\phi = 1, \theta = 2$ for the u and v nutrient yields respectively, and set the growth state size to $n = 10$ in our model parameters.

The nutrient uptake rate (k_u and k_v in our model), assuming the rate-limiting process occurs at the level of the cell membrane, can be estimated from a collision

frequency based calculation and primarily depend on the radius of the cell and the diffusion constant of the nutrient. Following Button [154], the specific affinity, $a_{\max}$, for a population of perfectly absorbing spherical cells with radius r_x can be written as

$$a_{\max} = \frac{10DM}{r_x^2} \text{ liters (g of cells)}^{-1}h^{-1} \tag{5.2}$$

where D is the diffusion constant of the nutrient and M its molecular weight, assuming a cell density of 1.08 g cm $^{-3}$ (approximately true for bacteria and yeast cells). To further comply with our model, we need to consider only collisions with the appropriate transporters. The reduction in specific affinity is related to the number of area of the cell regions occupied by the appropriate transporters, and is given by

$$a_s = a_{\max}\zeta = a_{\max}\left(\frac{N\pi r_s^2}{4\pi r_x^2}\right) \tag{5.3}$$

where the theoretical absorption coefficient, ζ, depends on the number of transporters, N, and the radius of the effective surface area of the transporter, r_s. Our genetic manipulations of the transporter concentrations are targeting the ζ — which acts as the parametrizable nutrient preference phenotype, s_x, in our model — by changing the number of transporters, N. The constructs would, ideally, span the interval $\zeta \in [0, 1]$ for each nutrient in our setup; in practice, however, this may be difficult to achieve. According to our model of constrained nutrient preference, we will need to construct the ζ such that $\zeta_u + \zeta_v =$ constant, ideally 1 for each strain in our discretized phenotype space. The a_s for each construct can be measured experimentally, following Button [154], from single-nutrient batch growth experiments and assuming Monod growth kinetics.

We are left with the metabolic rate of each nutrient (k_ϕ, k_θ), which is harder to estimate because it involves many enzymatic reactions. It may be possible to use published results from flux balance analysis of well-characterized metabolic pathways [95] to effective (single-step) metabolic rate. However, such results are typically not available for all of the nutrients that we may want to use. Since we can estimate the other nutrient parameters, and the chemostat parameters can be controlled, we can estimate the nutrient metabolic rate constant by fitting a single-phenotype single-nutrient growth experiment, recording the concentration

of cells in time (e.g. by optical density), to our model. We will assume that the metabolic rate of the nutrient is not dependent on the metabolism of the second nutrient.

5.3.2 Competition experiments

Following characterization of the nutrient parameters, we are then ready to conduct the competition experiments between strains in our discretized phenotype space. Since our goal is to obtain an adaptation graph, we can achieve this by following the dynamic programming algorithm described in Section **??**, Figure **??**, proceeding in batches at each of the tree (co-residency) level. Each competition experiment will need to be carried out in replicate since the invasion probability of even a conditionally-more-fit phenotype from small cell numbers will be less than one (Section **??**).

For the first co-residency level, where a phenotypic strain ℓ_x 'invades' an environment with no resident, we proceed as follows. We first allow the chemostat, without cells, to equilibrate. We then inoculate with a small sample of phenotypic strain ℓ_x and allow the system to reach steady state. We can establish when the system has reached steady state by periodically sampling the chemostat outflow and measuring the optical density (OD) of the sample, which should plateau at steady state. We retain a sample from steady state for sequencing, together with the steady state OD. This process is repeated, in replicate, for every lineage in our discretized phenotype space. We can use these results to construct the first level of the adaptation map.

For the second co-residency level, we require all pairwise strain competitions, $\{\ell_y, \ell_x\}$, for each of the ℓ_x that can survive in the chemostat environment. After growing the ℓ_x to steady state, we introduce a small sample of the 'mutant' lineage ℓ_y. It is now more difficult to determine when steady state is reached from OD measurements since these do not discriminate between strains; however, our simulations suggest that the total population size should change, at least during invasion, and usually settles to a new population size (hence a new OD value) at steady state, relative to the initial condition. At steady state, then, we record the OD of the final, extant, population and keep a sample for sequencing.

When we have exhausted all pairwise competitions, we can determine the out-

come of each competition in a single, multiplexed sequencing run. The fraction of replicate experiments where ℓ_y invades ℓ_x gives the empirical probability of surviving accidental extinction, $P_s(r_y|r_{\boldsymbol{x}}, \boldsymbol{x})$, in Equation **??**. $P_I(r_y|r_{\boldsymbol{x}}, \boldsymbol{x})$ is zero if no replicate resulted in invasion, or one if at least one replicate did. These data will allow us to construct the second level of the adaptation map.

We then repeat the process for the third co-residency level, where a strain ℓ_z invades all possible combinations of $\{\ell_x, \ell_y\}$ that can co-exist. Because we know from our sequencing results which co-existence combinations are viable, we only need to do those experiments where ℓ_z invades a co-existing pair of $\{\ell_x, \ell_y\}$. The experiments are carried out as before, with the batch sequencing reaction at the end. Our theoretical analysis did not extend beyond a third co-residency level because a triple co-existence was not possible in our model; therefore, we anticipate that the experimental adaptation map will also terminate at the same level.

5.3.3 Comparing data with model predictions

At the end of the experiment, we should have sufficient data to construct an adaptation map like the one shown in Figure **??**. At this stage, the adaptation map predicted by the model should be directly comparable to the one obtained experimentally. As a first pass, we can test whether the map edges exist – i.e. comparing the 'deterministic' invasion probabilities, $P_I(r_y|r_{\boldsymbol{x}}, \boldsymbol{x})$; for the cases where it is possible that a mutant phenotype invades ($P_I \neq 0$), we can also compare the probability of surviving accidental extinction, $P_s(r_y|r_{\boldsymbol{x}}, \boldsymbol{x})$, with the experimental fraction of replicate experiments where ℓ_y invaded.

In principle, agreement of the predicted and empirical adaptation maps implies that any adaptation graph constructed by parsing the maps with a given mutation kernel should yield identical graphs. A corollary of this statement is that the long-term stationary behavior in experiments will agree with the theoretically predicted one. Because we have constructed all parametrized phenotypes, the mutation kernel is under the experimenter's control. For a naturally-adapting system, however, it will usually be harder to know the phenotypic mutation kernel. However, notable exceptions might exist at the level of a single gene (for example, one conferring an antibiotic resistant phenotype [19]), where the effect of single-nucleotide mutations can be quantified; at the organismal level, a labora-

tory evolution experiment of antibiotic resistance in bacteria [155] has been used to construct a genotype-phenotype map that can be used as a coarse mutation kernel.

The parameter exploration we conducted for the two-nutrient model (Section **??**) can be tested experimentally. To do so, we would need to test multiple chemostat parameter sets: either constructing the entire adaptation map, as described earlier, or only verifying that the stationary phenotype composition predicted by the model cannot be invaded by any phenotypic strain to test the stationary behavior. In our analysis, we found that the model parameters with the largest impact on (the marginal probability of) the long-term stationary behavior of adaptation were the nutrient concentration ratios and the chemostat dilution rate (Figure 4.3.10). Since these are easily controlled in the experimental setup, compared to finding alternative nutrient sources with novel parameters, we do not need to make new constructs to explore the regions of parameter space where we expect to see considerable variation in stationary behavior.

5.4 Future directions

New research directions that merit further investigation emerged as a result of this work. Here we highlight some of the most promising ones:

- While our analytical decomposition of the alternative metabolic pathways model holds up to numerical examination, we have not been able to produce a rigorous proof; as such, the universality of the expression combining the independent subsystems, which correspond to ergodic trajectories through the growth state space, remains a conjecture. Second, as we noted in Section **??**, it is tempting to speculate that an even better transformation of the independent subsystems exists that can map their combination to the complete model with alternative pathways for every part of the dynamical trajectory - not just the steady state. Finally, it may be possible to lift the restrictions that we placed on the growth trajectories for the decomposition to work: namely, that a single state (the first state) acts as the single point of re-entry into the population structure after a cell has replicated. Such an extension would allow us to derive similar decompositions (of a phenotype's

overall fitness) to the fitnesses of cells following independent trajectories for more complicated models, including those with multiple nutrients.

- One limitation of our approach is the assumption that the appearance of novel (mutant) phenotypes is well-separated in time from other mutation events. This is an inherent simplifying assumption in adaptive dynamics theory [76] that we have initially opted to allow. As discussed in the introduction **??**, however, clonal interference, whereby many beneficial mutations can transiently co-exist - and compete - out of equilibrium, may have important implications for adaptive evolution [29, 30, 31].

  Standard adaptive dynamics theory is not yet equipped to rigorously handle such transient diversity. Nevertheless, our computational approach can be extended to inject mutations prior to a model's dynamical steady state according to a mutation kernel. From other work (not described in this thesis), using approximate stochastic methods to model rare events [156], we have some evidence to believe that the effects of transient phenotypic diversity on adaptive evolution can be significant under certain conditions. As such, this and related avenues should yield interesting insights on the impact of clonal interference.

- There is also much potential for improving the methods for analyzing the data generated from adaptation maps and graphs. At the level of an individual graph, for example, it should be possible to use modern concepts from social and information network analysis to locate vertices - i.e phenotypic compositions - that are important for the adaptation process in different ways (see, for example, Appendix A).

  More work can also be done at the level of graph ensembles, like those generated during random sampling of the parameter space. In Sections **??** and **??**, we manually classified the graphs according to their stationary behavior. We then found transformations that map the parameter space to feature spaces that separate the classes and trained models to classify parameter sets. This approach (or at least certain aspects of it) can be automated and included in the software, although more robust machine learning techniques may be required.

- In Section **??** we described an alternative method for producing adaptation graphs. This method first relies on continuous sampling of the adaptation

parameter space to produce trajectories and only subsequently discretizes the adaptation space to construct the adaptation graph/process. By removing the restriction of discrete phenotypes, barriers to the phenotypic trajectory that may have otherwise affected the adaptation process are removed. At the same time, however, the assumption of continuous phenotypes is not always appropriate.

There are still technical concerns to be addressed in the implementation of this method. For example, determining a 'best' way to discretize the parameters realized during an adaptation trajectory has proven to be more difficult, and more ambiguous, than originally anticipated. The method is also more efficient, because the system will naturally spend more time in the more 'interesting' areas of the parameter space and less time in areas not likely to be encountered during the adaptation process. This is both good and bad: on the one hand, we can more accurately represent (through finer post-trajectory discretization) the important parameter regions; on the other hand, we may require many trajectories (possibly from many initial conditions) before we encounter the more uncommon parameter regions that the process may explore.

Nevertheless, the alternative method's computational merits may outweigh the issues. The original mapping method simulates a number of models that grows as the power of the number of co-residents, which becomes an issue for models that can support extensive lineage co-existence. Further, it becomes even more computationally prohibitive to simulate many out-of-equilibrium mutant invasions using the mapping method (if we wish to investigate clonal interference, for example), whereas the alternative method can more naturally accommodate such events.

- Phenotypic constraints and resource allocation trade-offs were originally proposed in the context of evolutionary theories of ageing [88]. In other work not described in this thesis we have applied similar methods to investigate the adaptive evolution of replicative lifespan (the number of daughters a cell can produce) as a consequence of limited resource allocation between cellular damage repair, biomass production, and nutrient use. In these models, the resource allocation strategy determines the lineage's survival curve. It should be possible to fit experimentally obtained survival curves [157] to theoretical survival curves generated by the adaptively stable phenotype

of a given set of model parameters. By finding the model parameters that generate the observed survival curve, which we may assume is adaptively stable, it may be possible to gain some insight into the environmental pressures that drove the evolution of cellular ageing for various microbial species and strains, and also possible directions in which replicative lifespan may change as a consequence of changing environmental conditions.

5.5 Closing remarks

Adaptive evolution is central to many areas of biology that affect human and non-human systems. It is the process by which pathogenicity emerges, antibiotic resistance increases, and cancer progresses. In this work we have studied some aspects of adaptive evolution, focusing on developing methods that explore the scope of adaptation that we may encounter in nature. We have applied our work to established and new models of adaptive dynamics, motivated by understanding the effects of phenotype constraints on adaptation. In doing so, we have affirmed prior knowledge, uncovered new possibilities in which adaptation may manifest, and demonstrated the need for analytical and numerical methods that expand the way in which adaptive dynamics are modeled and studied. We hope that our own methodology can serve as a useful utility for analyzing new and existing models of adaptation, and also serve as a starting point from which new methods can be developed.

Appendices

Appendix A

Three substitutable nutrients: graph centrality measures

The average graphs from the $1,000$ adaptation maps are re-interpreted by highlighting (resizing) each vertex according to its centrality value. Three centrality metrics were used:

1. Vertex degree: the number of incoming plus outgoing edges. This metrics identifies vertices that influence other vertices in their immediate neighborhood, for example by accepting many edges (a strongly adaptively stable phenotype composition) and/or by having many outgoing edges (a weakly adaptively stable phenotype composition).

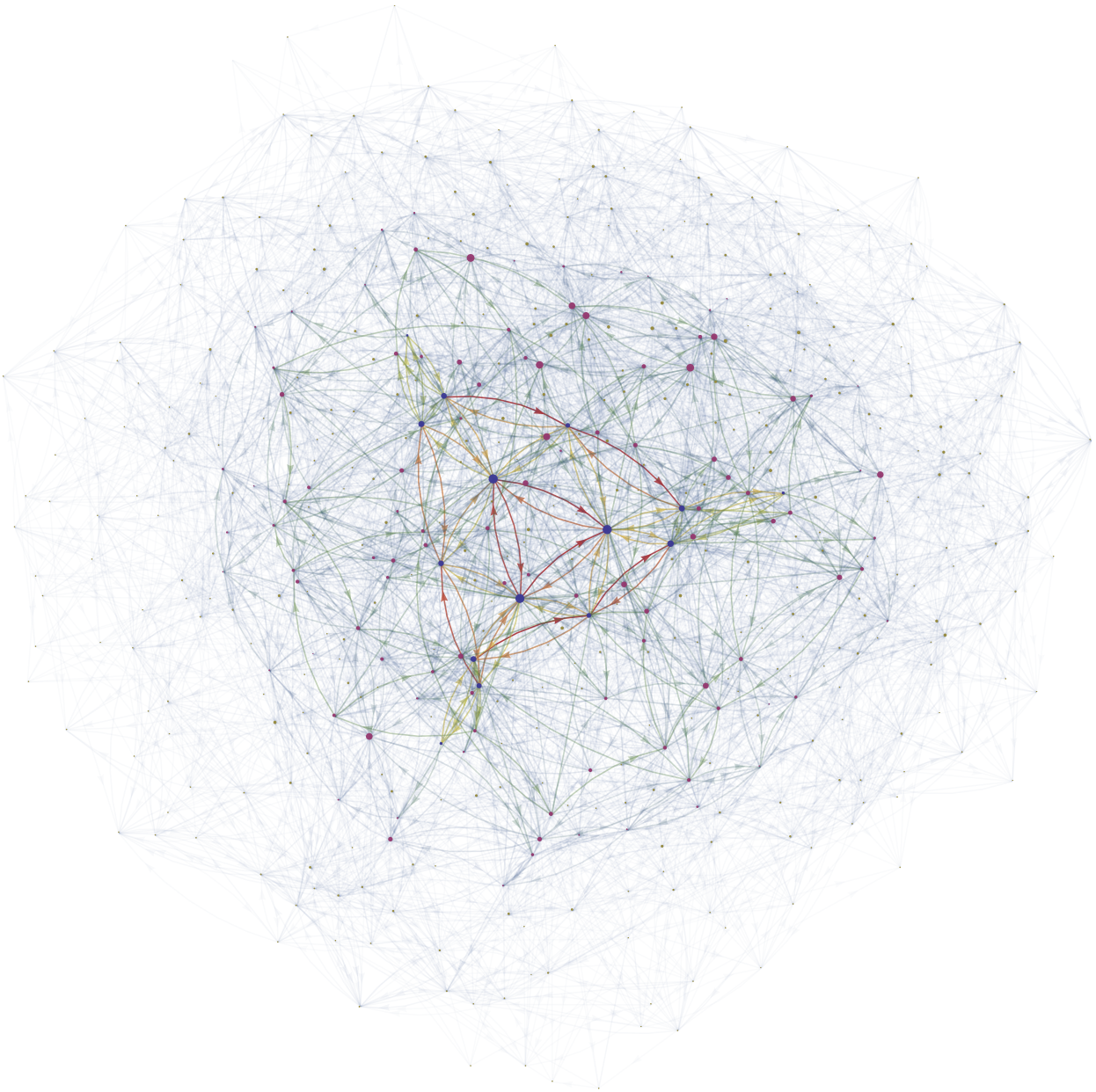
2. Vertex betweenness: a normalized measure of the number of shortest paths through the vertex.

$$g(v) = \sum_{s \neq v \neq t} \frac{\sigma_{s,t}(v)}{\sigma_{s,t}}$$

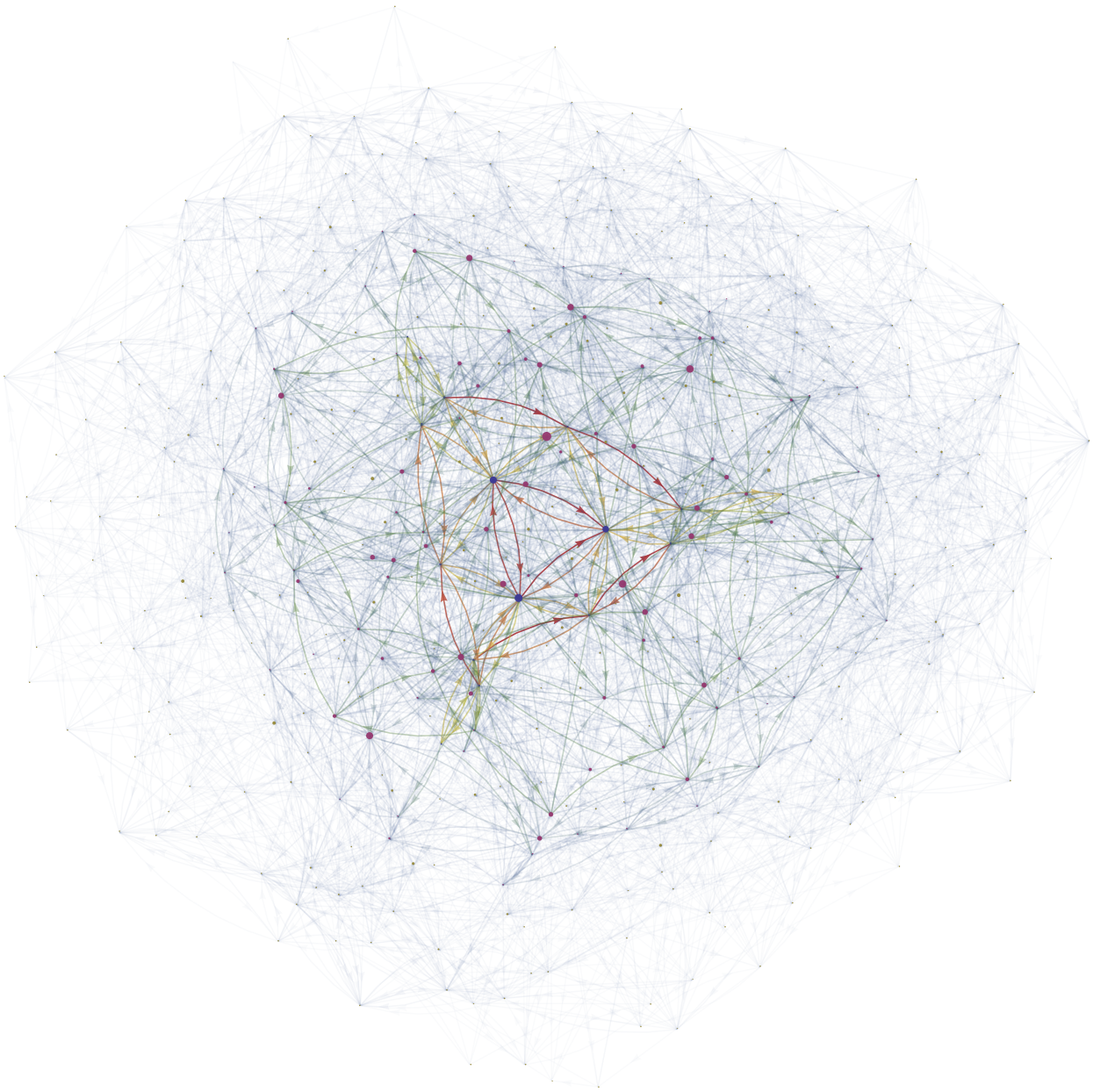
   where $\sigma_{s,t}(v)$ is the number of shortest paths from vertex s to vertex t through vertex v and $\sigma_{s,t}$ is the total number of shortest paths from s to t. This metric identifies vertices that appear frequently on shortest paths between two other randomly chosen vertices; as such, they represent phenotype compositions that frequently acts as bridges as the adaptation process approaches stationarity.

3. Vertex closeness: the reciprocal of the average distance from the vertex to all other vertices connected to it. This metric identifies vertices that

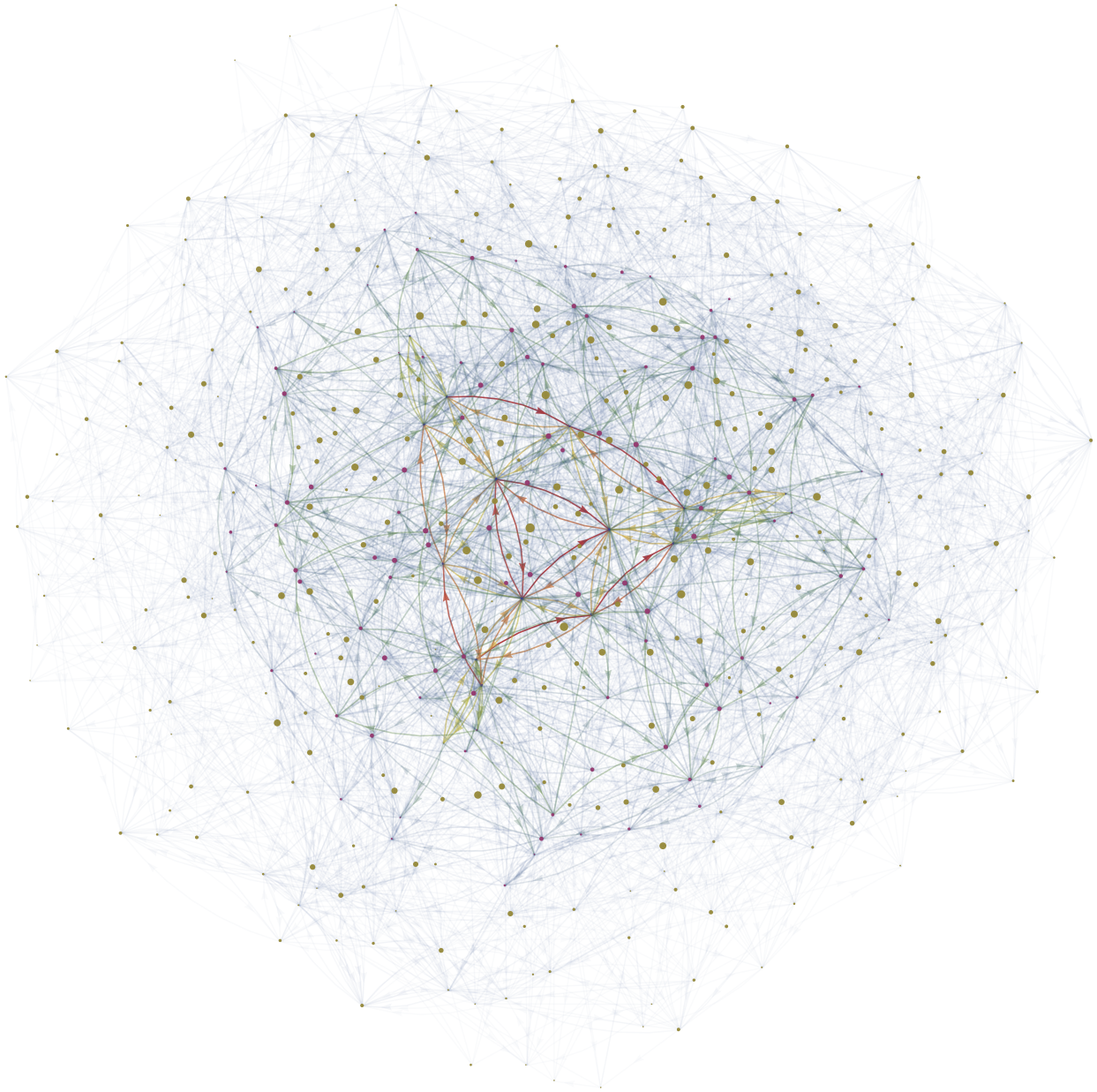
influence the entire graph and are important for providing the adaptation process with quick paths to stationary vertices.



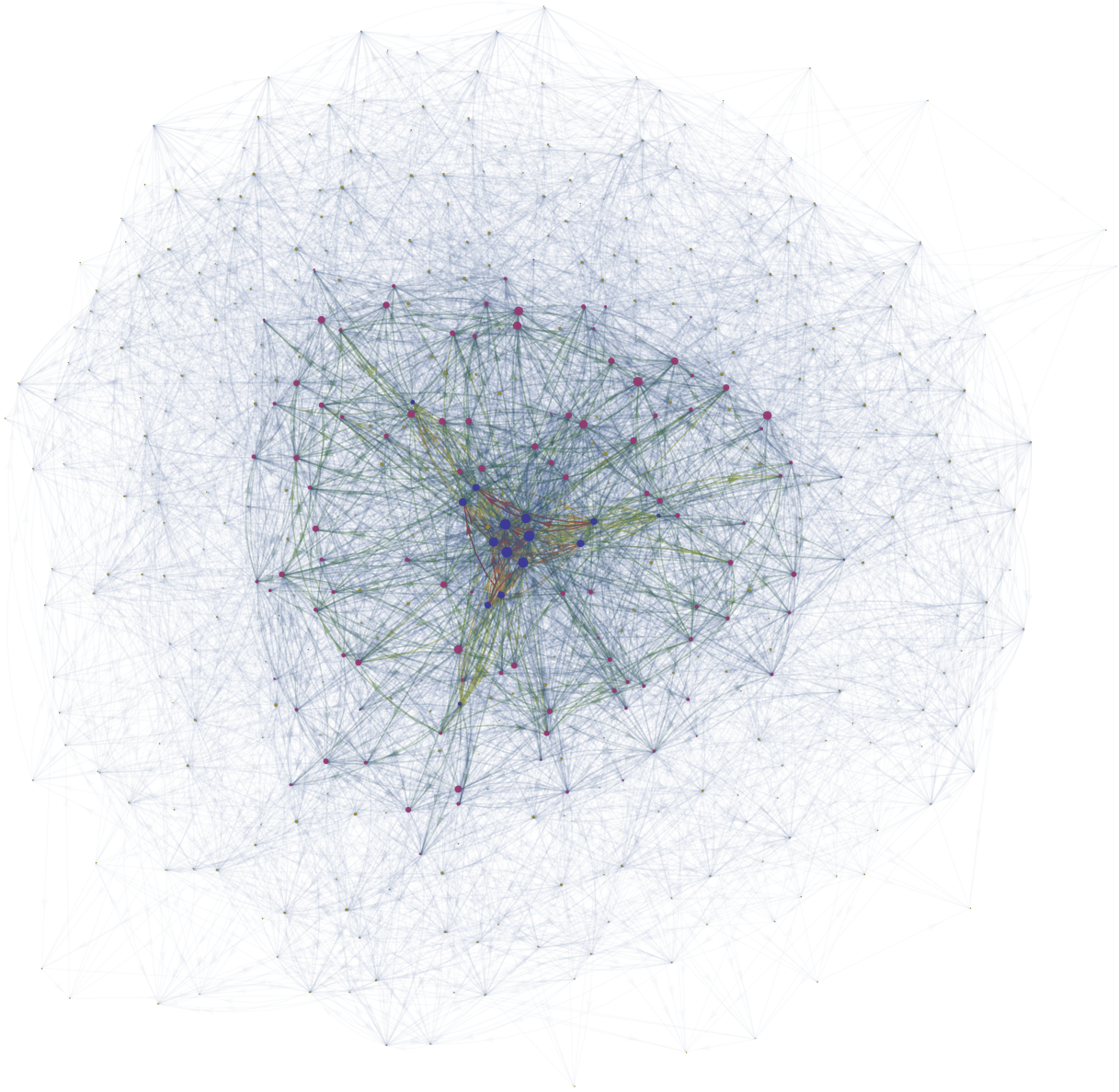
Centrality = vertex degree ($\mu_{max} = \sqrt{2}/4$).



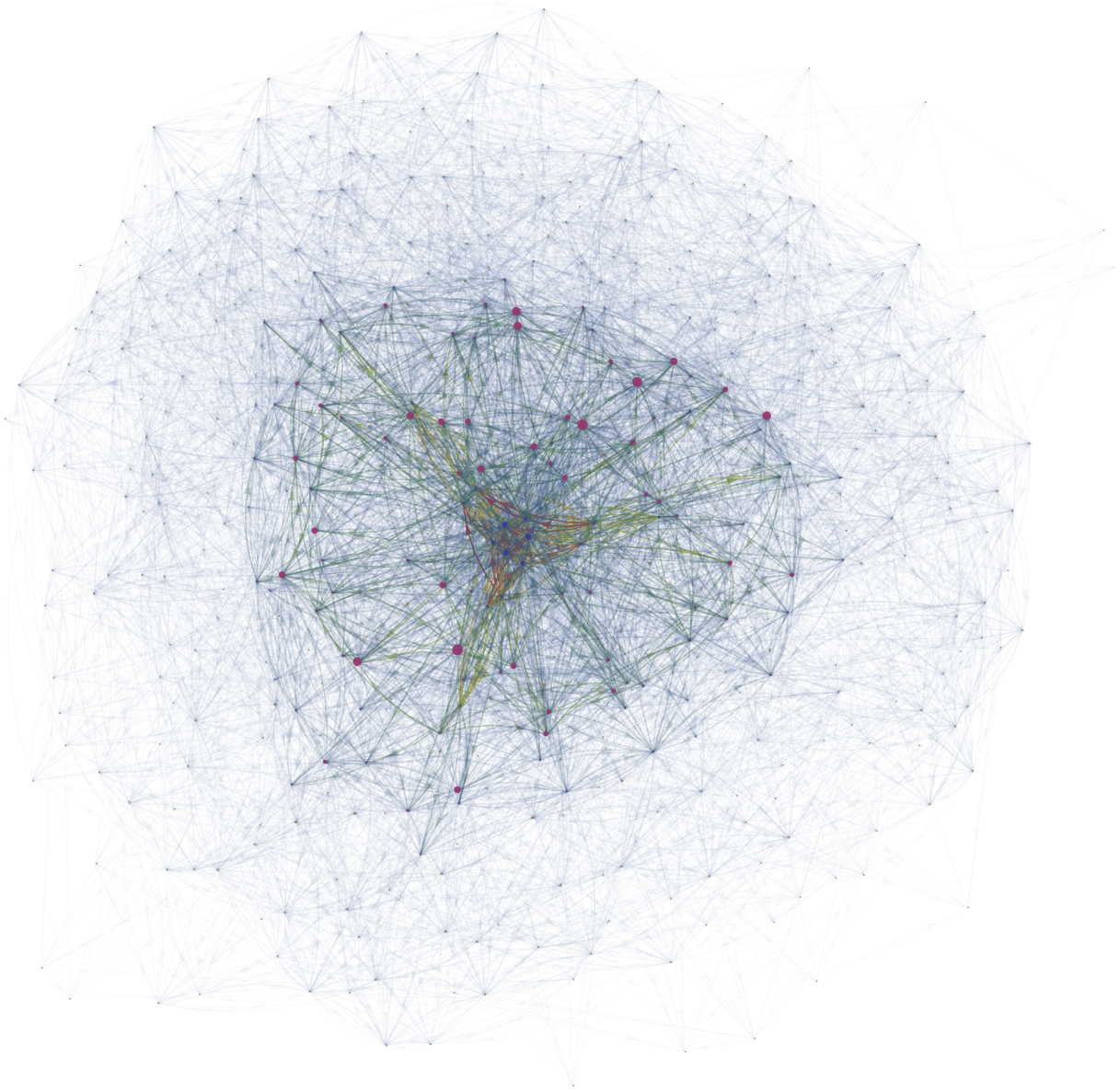
Centrality = vertex betweenness ($\mu_{max} = \sqrt{2}/4$).



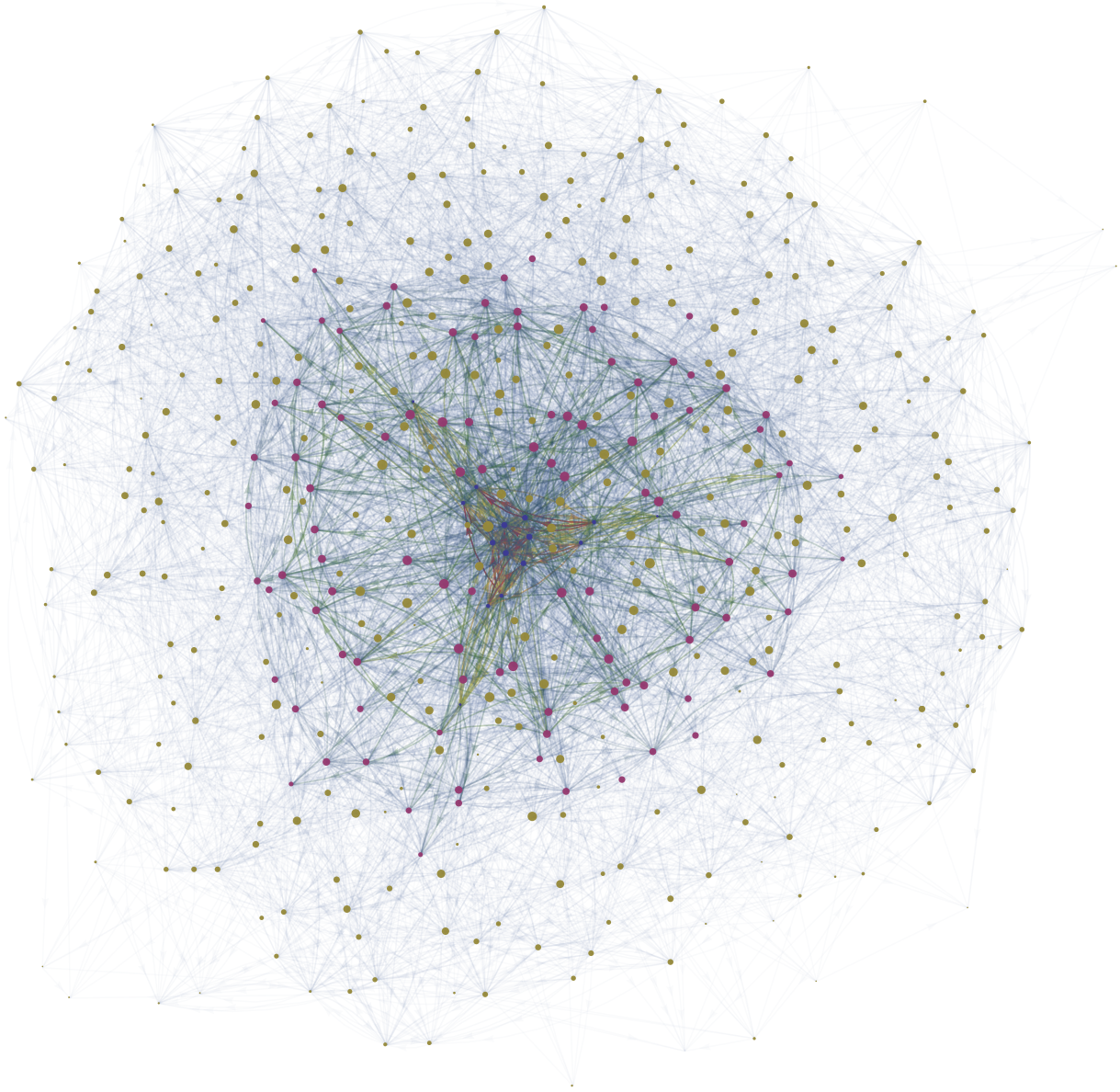
Centrality = vertex closeness ($\mu_{max} = \sqrt{2}/4$).



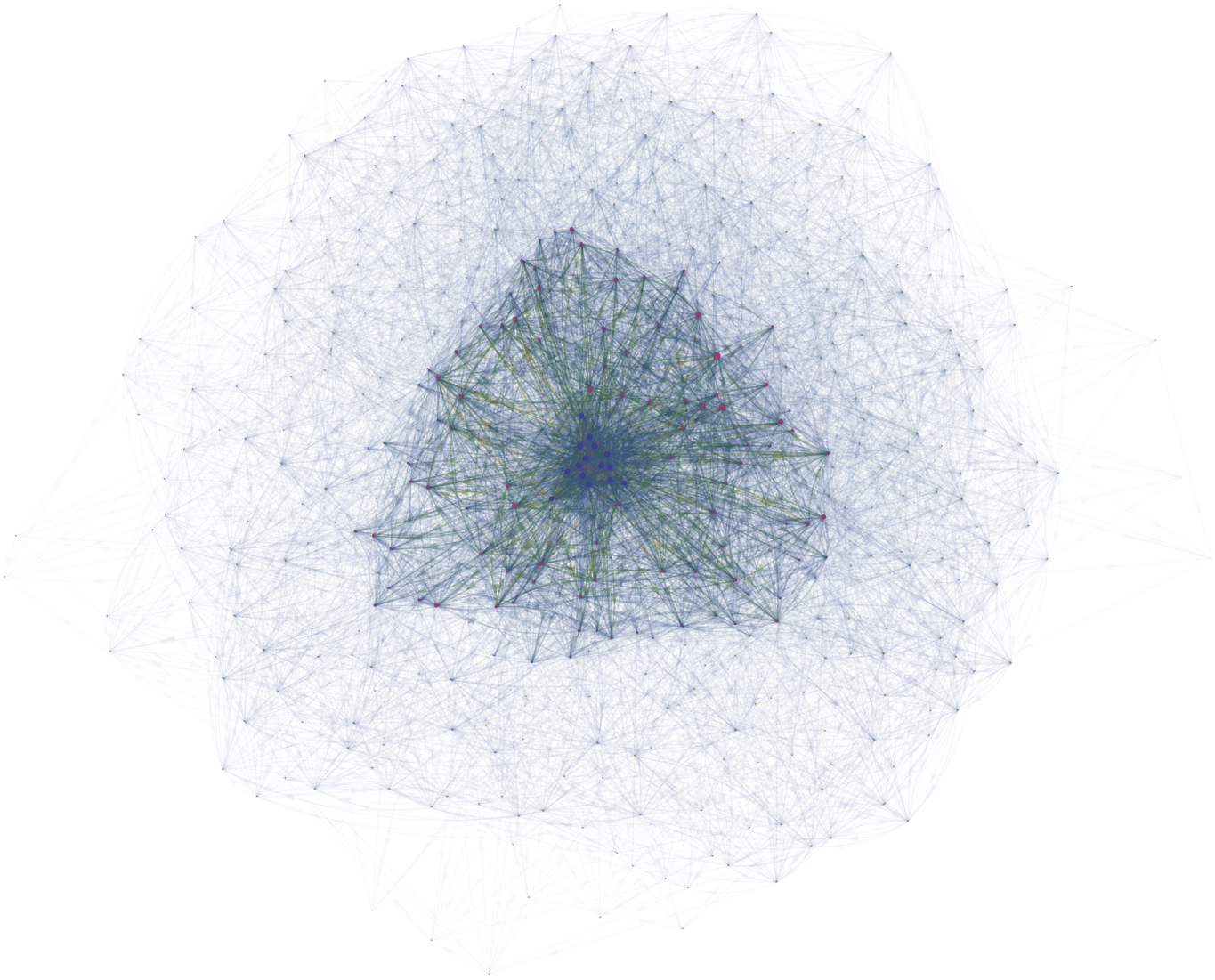
Centrality = vertex degree ($\mu_{max} = \sqrt{2}/2$).



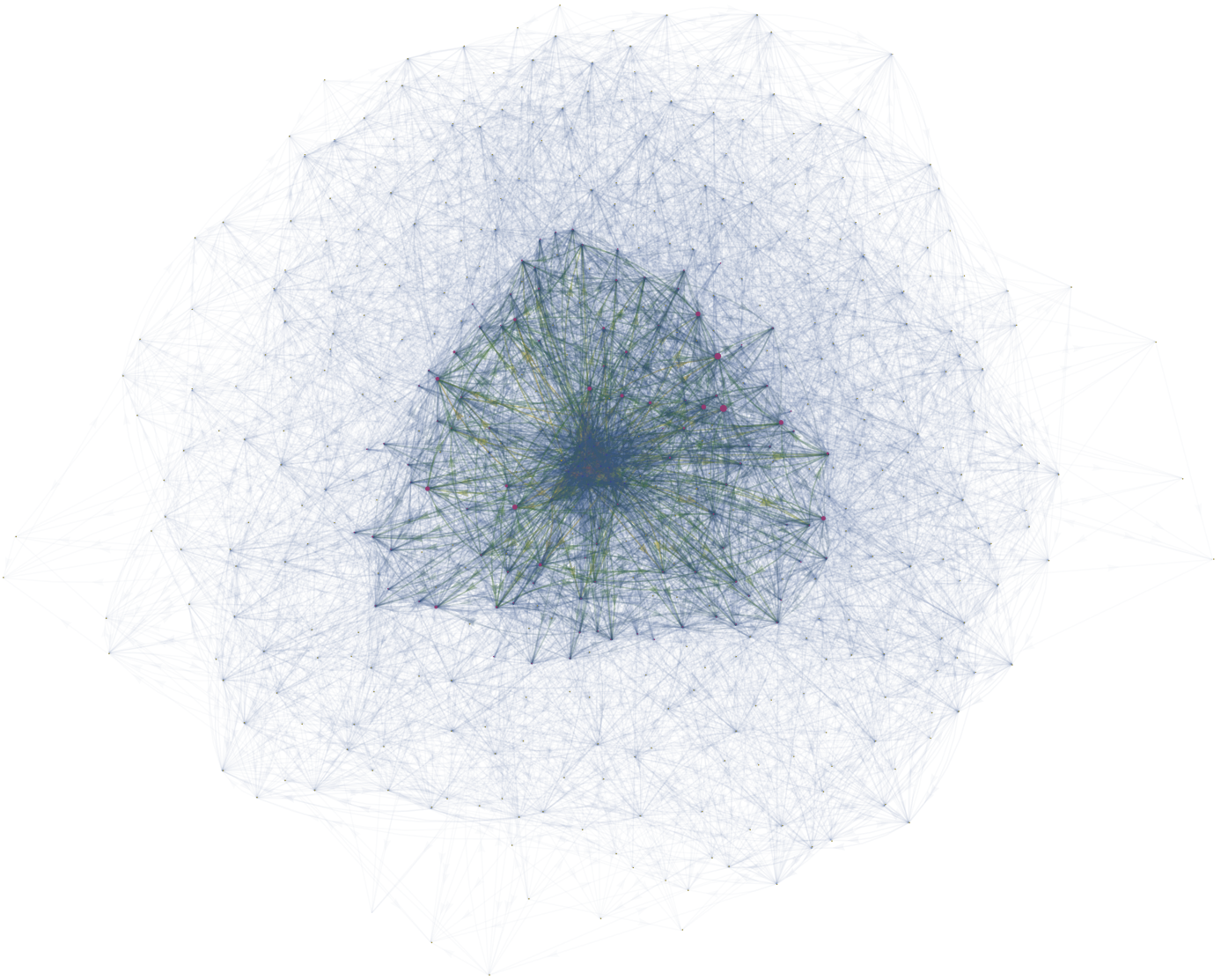
Centrality = vertex betweenness ($\mu_{max} = \sqrt{2}/2$).



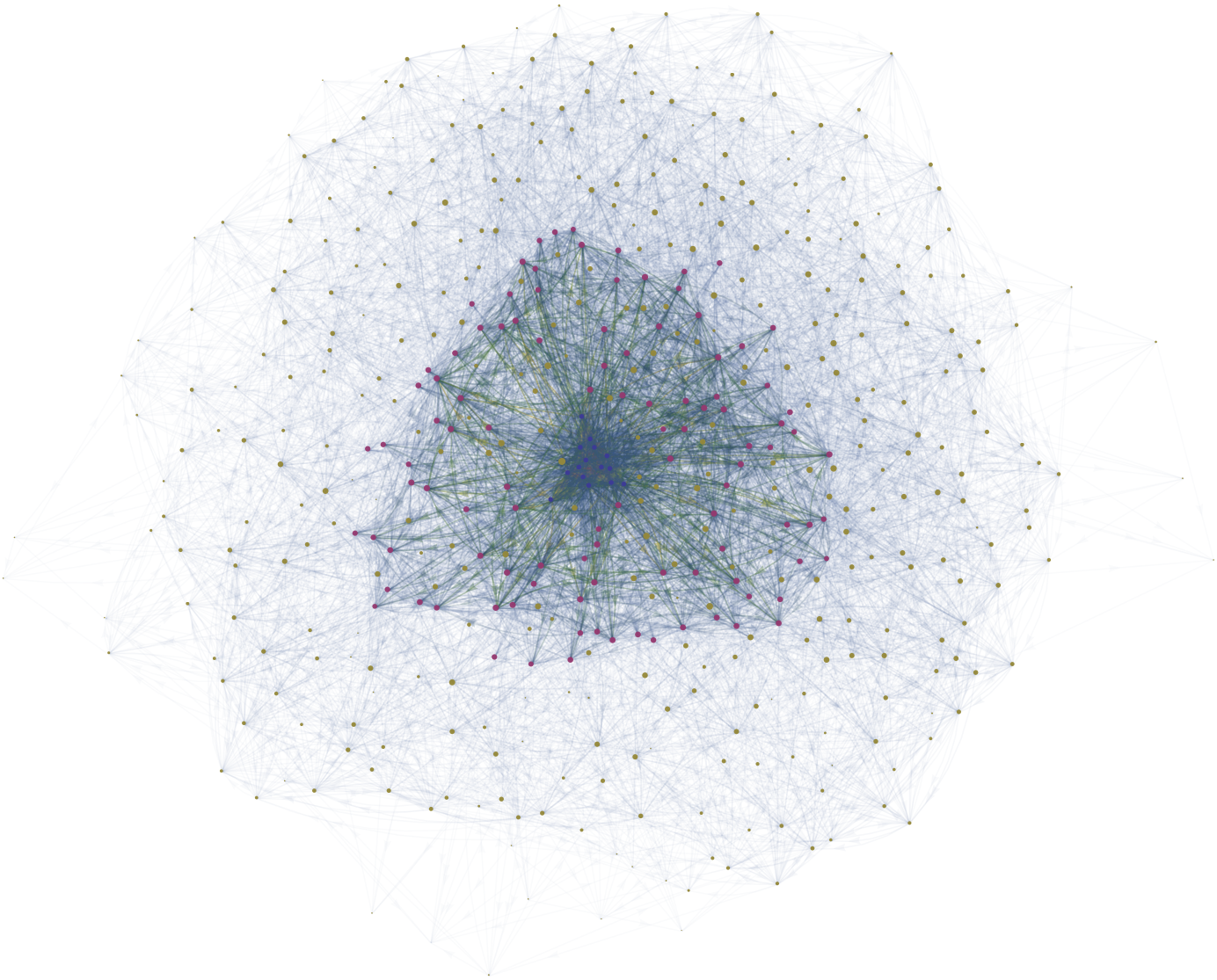
Centrality = vertex closeness ($\mu_{max} = \sqrt{2}/2$).



Centrality = vertex degree ($\mu_{max} = \sqrt{2}$).



Centrality = vertex betweenness ($\mu_{max} = \sqrt{2}$).



Centrality = vertex closeness ($\mu_{max} = \sqrt{2}$).

Bibliography

[1] B. Charlesworth, "The good fairy godmother of evolutionary genetics," *Current Biology*, vol. 6, no. 3, p. 220, 1996.

[2] M. Kimura, "On the probability of fixation of mutant genes in a population," *Genetics*, vol. 47, no. 6, p. 713, 1962.

[3] M. Kimura, "Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution," 1977.

[4] M. Kimura, *The neutral theory of molecular evolution*. Cambridge University Press, 1984.

[5] J. E. Barrick and R. E. Lenski, "Genome dynamics during experimental evolution," *Nature Reviews Genetics*, vol. 14, no. 12, pp. 827–839, 2013.

[6] P. A. P. Moran, "Random processes in genetics," vol. 54, no. 01, pp. 60–71, 1958.

[7] S. P. Otto and M. C. Whitlock, "Fixation probabilities and times," *eLS*, 2001.

[8] R. Kassen and T. Bataillon, "Distribution of fitness effects among beneficial mutations before selection in experimental populations of bacteria," *Nature genetics*, vol. 38, no. 4, pp. 484–488, 2006.

[9] H. A. Orr, "The population genetics of adaptation: the adaptation of dna sequences," *Evolution*, vol. 56, no. 7, pp. 1317–1330, 2002.

[10] A. Sancar, L. A. Lindsey-Boltz, K. Ünsal-Kaçmaz, and S. Linn, "Molecular mechanisms of mammalian dna repair and the dna damage checkpoints," *Annual review of biochemistry*, vol. 73, no. 1, pp. 39–85, 2004.

[11] C. D. Schlichting, M. Pigliucci, *et al.*, *Phenotypic evolution: a reaction norm perspective*. Sinauer Associates Incorporated, 1998.

[12] R. Lande, "Natural selection and random genetic drift in phenotypic evolution," *Evolution*, pp. 314–334, 1976.

[13] M. Nei, "The new mutation theory of phenotypic evolution," *Proceedings of the National Academy of Sciences*, vol. 104, no. 30, pp. 12235–12242, 2007.

[14] R. C. Lewontin, "A general method for investigating the equilibrium of gene frequency in a population," *Genetics*, vol. 43, no. 3, p. 419, 1958.

[15] R. A. Fisher, "The genetical theory of natural selection: a complete variorum edition," 1930.

[16] R. Olendorf, F. H. Rodd, D. Punzalan, A. E. Houde, C. Hurt, D. N. Reznick, and K. A. Hughes, "Frequency-dependent survival in natural guppy populations," *Nature*, vol. 441, no. 7093, pp. 633–636, 2006.

[17] E. Trachtenberg, B. Korber, C. Sollars, T. B. Kepler, P. T. Hraber, E. Hayes, R. Funkhouser, M. Fugate, J. Theiler, Y. S. Hsu, *et al.*, "Advantage of rare hla supertype in hiv disease progression," *Nature medicine*, vol. 9, no. 7, pp. 928–935, 2003.

[18] M. Le Gac, J. Plucain, T. Hindré, R. E. Lenski, and D. Schneider, "Ecological and evolutionary dynamics of coexisting lineages during a long-term experiment with escherichia coli," *Proceedings of the National Academy of Sciences*, vol. 109, no. 24, pp. 9487–9492, 2012.

[19] D. M. Weinreich, N. F. Delaney, M. A. DePristo, and D. L. Hartl, "Darwinian evolution can follow only very few mutational paths to fitter proteins," *science*, vol. 312, no. 5770, pp. 111–114, 2006.

[20] D. I. Rosenbloom, A. L. Hill, S. A. Rabi, R. F. Siliciano, and M. A. Nowak, "Antiretroviral dynamics determines hiv evolution and predicts therapy outcome," *Nature medicine*, vol. 18, no. 9, pp. 1378–1385, 2012.

[21] J. A. Engelman, L. Chen, X. Tan, K. Crosby, A. R. Guimaraes, R. Upadhyay, M. Maira, K. McNamara, S. A. Perera, Y. Song, *et al.*, "Effective use of pi3k and mek inhibitors to treat mutant kras g12d and pik3ca h1047r murine lung cancers," *Nature medicine*, vol. 14, no. 12, pp. 1351–1356, 2008.

[22] E. D. Pleasance, R. K. Cheetham, P. J. Stephens, D. J. McBride, S. J. Humphray, C. D. Greenman, I. Varela, M.-L. Lin, G. R. Ordóñez, G. R. Bignell, *et al.*, "A comprehensive catalogue of somatic mutations from a human cancer genome," *Nature*, vol. 463, no. 7278, pp. 191–196, 2010.

[23] S. E. Luria and M. Delbrück, "Mutations of bacteria from virus sensitivity to virus resistance," *Genetics*, vol. 28, no. 6, p. 491, 1943.

[24] A. Novick and L. Szilard, "Experiments with the chemostat on spontaneous mutations of bacteria," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 36, no. 12, p. 708, 1950.

[25] D. E. Dykhuizen and D. L. Hartl, "Selection in chemostats.," *Microbiological Reviews*, vol. 47, no. 2, p. 150, 1983.

[26] P. A. Lind and D. I. Andersson, "Whole-genome mutational biases in bacteria," *Proceedings of the National Academy of Sciences*, vol. 105, no. 46, pp. 17878–17883, 2008.

[27] H. Lee, E. Popodi, H. Tang, and P. L. Foster, "Rate and molecular spectrum of spontaneous mutations in the bacterium escherichia coli as determined by whole-genome sequencing," *Proceedings of the National Academy of Sciences*, p. 201210309, 2012.

[28] K. Atwood, L. K. Schneider, and F. J. Ryan, "Periodic selection in escherichia coli," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 37, no. 3, p. 146, 1951.

[29] S. F. Levy, J. R. Blundell, S. Venkataram, D. A. Petrov, D. S. Fisher, and G. Sherlock, "Quantitative evolutionary dynamics using high-resolution lineage tracking," *Nature*, 2015.

[30] J. R. Blundell and S. F. Levy, "Beyond genome sequencing: Lineage tracking with barcodes to study the dynamics of evolution, infection, and cancer," *Genomics*, vol. 104, no. 6, pp. 417–430, 2014.

[31] G. I. Lang, D. P. Rice, M. J. Hickman, E. Sodergren, G. M. Weinstock, D. Botstein, and M. M. Desai, "Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations," *Nature*, vol. 500, no. 7464, pp. 571–574, 2013.

[32] C. C. Spencer, J. Tyerman, M. Bertrand, and M. Doebeli, "Adaptation increases the likelihood of diversification in an experimental bacterial lineage," *Proceedings of the National Academy of Sciences*, vol. 105, no. 5, pp. 1585–1589, 2008.

[33] M. Le Gac and M. Doebeli, "Epistasis and frequency dependence influence the fitness of an adaptive mutation in a diversifying lineage," *Molecular ecology*, vol. 19, no. 12, pp. 2430–2438, 2010.

[34] F. J. Poelwijk, D. J. Kiviet, D. M. Weinreich, and S. J. Tans, "Empirical fitness landscapes reveal accessible evolutionary paths," *Nature*, vol. 445, no. 7126, pp. 383–386, 2007.

[35] M. D. Herron and M. Doebeli, "Parallel evolutionary dynamics of adaptive diversification in escherichia coli," *PLoS biology*, vol. 11, no. 2, p. e1001490, 2013.

[36] J. Hofbauer and K. Sigmund, *Evolutionary games and population dynamics*. Cambridge University Press, 1998.

[37] D. Schluter, *The ecology of adaptive radiation*. Oxford University Press, 2000.

[38] A. Novick and L. Szilard, "Description of the chemostat," *Science*, vol. 112, pp. 715–716, Dec 1950.

[39] F. Jacob and J. Monod, "Genetic regulatory mechanisms in the synthesis of proteins," *Journal of molecular biology*, vol. 3, no. 3, pp. 318–356, 1961.

[40] H. L. Smith, *The theory of the chemostat: dynamics of microbial competition*, vol. 13. Cambridge university press, 1995.

[41] S.-B. Hsu, S. Hubbell, and P. Waltman, "A mathematical theory for single-nutrient competition in continuous cultures of micro-organisms," *SIAM Journal on Applied Mathematics*, vol. 32, no. 2, pp. 366–383, 1977.

[42] D. Tilman, "Resources: a graphical-mechanistic approach to competition and predation," *American Naturalist*, pp. 362–393, 1980.

[43] R. Maharjan, S. Seeto, L. Notley-McRobb, and T. Ferenci, "Clonal adaptive radiation in a constant environment," *Science*, vol. 313, no. 5786, pp. 514–517, 2006.

[44] O. Smith, H. Shugart, R. O'Neill, R. Booth, and D. McNaught, "Resource competition and an analytical model of zooplankton feeding on phytoplankton," *American Naturalist*, pp. 571–591, 1975.

[45] P. Taylor and P. L. Williams, "Theoretical studies on the coexistence of competing species under continuous-flow conditions," *Canadian Journal of Microbiology*, vol. 21, no. 1, pp. 90–98, 1975.

[46] J. B. Wilson, E. Spijkerman, and J. Huisman, "Is there really insufficient support for tilmans r* concept? a comment on miller et al.," *The american naturalist*, vol. 169, no. 5, pp. 700–706, 2007.

[47] S. Hubbell *et al.*, "Single-nutrient microbial competition: qualitative agreement between experimental and theoretically forecast outcomes," *Science*, vol. 207, no. 4438, pp. 1491–1493, 1980.

[48] F. J. Ayala, "Experimental invalidation of the principle of competitive exclusion," 1969.

[49] J. P. Grime *et al.*, "Competitive exclusion in herbaceous vegetation.," *Nature, UK*, vol. 242, no. 5396, pp. 344–347, 1973.

[50] T. E. Miller, J. H. Burns, P. Munguia, E. L. Walters, J. M. Kneitel, P. M. Richards, N. Mouquet, and H. L. Buckley, "A critical review of twenty years use of the resource-ratio theory," *The American Naturalist*, vol. 165, no. 4, pp. 439–448, 2005.

[51] S. A. Levin, "Multiple scales and the maintenance of biodiversity," *Ecosystems*, vol. 3, no. 6, pp. 498–506, 2000.

[52] R. McGehee and R. A. Armstrong, "Some mathematical problems concerning the ecological principle of competitive exclusion," *Journal of Differential Equations*, vol. 23, no. 1, pp. 30–52, 1977.

[53] R. M. May and R. H. Mac Arthur, "Niche overlap as a function of environmental variability," *Proceedings of the National Academy of Sciences*, vol. 69, no. 5, pp. 1109–1113, 1972.

[54] P. L. Chesson and T. J. Case, "Overview: nonequilibrium community theories: chance, variability, history," *Community ecology. Edited by J. Dia-*

mond and TJ Case. Harper and Row Publishers, Inc., New York, pp. 229–239, 1986.

[55] S. P. Hubbell and R. B. Foster, "Biology, chance, and history and the structure of tropical rain forest tree communities,"

[56] F. M. Stewart and B. R. Levin, "Partitioning of resources and the outcome of interspecific competition: a model and some general considerations," *American Naturalist*, pp. 171–198, 1973.

[57] A. L. Koch, "Competitive coexistence of two predators utilizing the same prey under constant environmental conditions," *Journal of Theoretical Biology*, vol. 44, no. 2, pp. 387–395, 1974.

[58] B. C. Baltzis and A. Fredrickson, "Coexistence of two microbial populations competing for a renewable resource in a non-predator-prey system," *Bulletin of mathematical biology*, vol. 46, no. 1, pp. 155–174, 1984.

[59] M. Eigen, J. McCaskill, and P. Schuster, "The molecular quasi-species," *Adv. Chem. Phys*, vol. 75, pp. 149–263, 1989.

[60] M. A. Nowak, "What is a quasispecies?," *Trends in Ecology & Evolution*, vol. 7, no. 4, pp. 118–121, 1992.

[61] X. Wei, S. K. Ghosh, M. E. Taylor, V. A. Johnson, E. A. Emini, P. Deutsch, J. D. Lifson, S. Bonhoeffer, M. A. Nowak, B. H. Hahn, *et al.*, "Viral dynamics in human immunodeficiency virus type 1 infection," *Nature*, vol. 373, no. 6510, pp. 117–122, 1995.

[62] M. A. Nowak, R. M. May, R. E. Phillips, S. Rowland-Jones, D. G. Lalloo, S. McAdam, P. Klenerman, B. Koppe, K. Sigmund, C. R. Bangham, *et al.*, "Antigenic oscillations and shifting immunodominance in hiv-1 infections," *Nature*, vol. 375, no. 6532, pp. 606–611, 1995.

[63] P. D. Taylor and L. B. Jonker, "Evolutionary stable strategies and game dynamics," *Mathematical biosciences*, vol. 40, no. 1, pp. 145–156, 1978.

[64] R. E. Beardmore, I. Gudelj, D. A. Lipson, and L. D. Hurst, "Metabolic trade-offs and the maintenance of the fittest and the flattest," *Nature*, vol. 472, no. 7343, pp. 342–346, 2011.

[65] J. M. Smith, "The theory of games and the evolution of animal conflicts," *Journal of theoretical biology*, vol. 47, no. 1, pp. 209–221, 1974.

[66] J. M. Smith, *Evolution and the Theory of Games*. Cambridge university press, 1982.

[67] M. A. Nowak and K. Sigmund, "Evolutionary dynamics of biological games," *science*, vol. 303, no. 5659, pp. 793–799, 2004.

[68] S. A. Geritz, G. Mesze, J. A. Metz, *et al.*, "Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree," *Evolutionary ecology*, vol. 12, no. 1, pp. 35–57, 1998.

[69] S. A. Geritz, J. A. Metz, É. Kisdi, and G. Meszéna, "Dynamics of adaptation and evolutionary branching," *Physical Review Letters*, vol. 78, no. 10, p. 2024, 1997.

[70] S. A. Geritz, É. Kisdi, G. Meszéna, and J. A. Metz, "Adaptive dynamics of speciation: ecological underpinnings," *Adaptive speciation. Cambridge University Press, Cambridge*, pp. 54–75, 2004.

[71] G. Meszéna, M. Gyllenberg, F. J. Jacobs, and J. A. Metz, "Link between population dynamics and dynamics of darwinian evolution," *Physical review letters*, vol. 95, no. 7, p. 078105, 2005.

[72] M. Durinx, J. H. Metz, and G. Meszéna, "Adaptive dynamics for physiologically structured population models," *Journal of mathematical biology*, vol. 56, no. 5, pp. 673–742, 2008.

[73] F. Dercole and S. Rinaldi, *Analysis of Evolutionary Processes: The Adaptive Dynamics Approach and Its Applications: The Adaptive Dynamics Approach and Its Applications*. Princeton University Press, 2008.

[74] É. Kisdi and S. A. Geritz, "Adaptive dynamics: a framework to model evolution in the ecological theatre," *Journal of mathematical biology*, vol. 61, no. 1, pp. 165–169, 2010.

[75] H. C. Ito and U. Dieckmann, "Evolutionary branching under slow directional evolution," *Journal of theoretical biology*, vol. 360, pp. 290–314, 2014.

[76] D. Waxman and S. Gavrilets, "20 questions on adaptive dynamics," *Journal of evolutionary biology*, vol. 18, no. 5, pp. 1139–1154, 2005.

[77] M. Kirkpatrick and F. Rousset, "Wright meets ad: not all landscapes are adaptive," *Journal of evolutionary biology*, vol. 18, no. 5, pp. 1166–1169, 2005.

[78] P. Abrams, "Adaptive dynamics vs.adaptive dynamics," *Journal of evolutionary biology*, vol. 18, no. 5, pp. 1162–1165, 2005.

[79] S. Gourbiere and J. Mallet, "Has adaptive dynamics contributed to the understanding of adaptive and sympatric speciation?," *Journal of evolutionary biology*, vol. 18, no. 5, pp. 1201–1204, 2005.

[80] H. Spencer and M. Feldman, "Adaptive dynamics, game theory and evolutionary population genetics," *Journal of evolutionary biology*, vol. 18, no. 5, pp. 1191–1193, 2005.

[81] H. Kokko, "Useful ways of being wrong," *Journal of evolutionary biology*, vol. 18, no. 5, pp. 1155–1157, 2005.

[82] T. Van Dooren, "The future of a mutation-limited tool-box," *Journal of evolutionary biology*, vol. 18, no. 5, pp. 1158–1161, 2005.

[83] M. Doebeli and U. Dieckmann, "Adaptive dynamics as a mathematical tool for studying the ecology of speciation processes," *Journal of evolutionary biology*, vol. 18, no. 5, pp. 1194–1200, 2005.

[84] S. A. Geritz and M. Gyllenberg, "Seven answers from adaptive dynamics," *Journal of evolutionary biology*, vol. 18, no. 5, pp. 1174–1177, 2005.

[85] J.-H. S. Hofmeyr and A. Cornish-Bowden, "Regulating the cellular economy of supply and demand," *FEBS letters*, vol. 476, no. 1, pp. 47–51, 2000.

[86] S. Hui, J. M. Silverman, S. S. Chen, D. W. Erickson, M. Basan, J. Wang, T. Hwa, and J. R. Williamson, "Quantitative proteomic analysis reveals a simple strategy of global resource allocation in bacteria," *Molecular systems biology*, vol. 11, no. 2, p. 784, 2015.

[87] C. N. Hinshelwood, "136. on the chemical kinetics of autosynthetic systems," *J. Chem. Soc.*, pp. 745–755, 1952.

[88] T. B. Kirkwood, "Evolution of ageing.," *Nature*, vol. 270, no. 5635, pp. 301–304, 1977.

[89] T. B. Kirkwood and R. Holliday, "The evolution of ageing and longevity," *Proceedings of the Royal Society of London. Series B. Biological Sciences*, vol. 205, no. 1161, pp. 531–546, 1979.

[90] T. Kirkwood, *Accuracy in Molecular Processes: Its Control and Relevance to Living System*. Springer Science & Business Media, 2012.

[91] T. B. Kirkwood and D. P. Shanley, "Food restriction, evolution and ageing," *Mechanisms of ageing and development*, vol. 126, no. 9, pp. 1011–1016, 2005.

[92] E. Dekel and U. Alon, "Optimality and evolutionary tuning of the expression level of a protein," *Nature*, vol. 436, no. 7050, pp. 588–592, 2005.

[93] A. Wagner, "Energy constraints on the evolution of gene expression," *Molecular biology and evolution*, vol. 22, no. 6, pp. 1365–1374, 2005.

[94] R. U. Ibarra, J. S. Edwards, and B. O. Palsson, "Escherichia coli k-12 undergoes adaptive evolution to achieve in silico predicted optimal growth," *Nature*, vol. 420, no. 6912, pp. 186–189, 2002.

[95] J. Edwards and B. Palsson, "The escherichia coli mg1655 in silico metabolic genotype: its definition, characteristics, and capabilities," *Proceedings of the National Academy of Sciences*, vol. 97, no. 10, pp. 5528–5533, 2000.

[96] O. Shoval, H. Sheftel, G. Shinar, Y. Hart, O. Ramote, A. Mayo, E. Dekel, K. Kavanagh, and U. Alon, "Evolutionary trade-offs, pareto optimality, and the geometry of phenotype space," *Science*, vol. 336, no. 6085, pp. 1157–1160, 2012.

[97] M. Scott, C. W. Gunderson, E. M. Mateescu, Z. Zhang, and T. Hwa, "Interdependence of cell growth and gene expression: origins and consequences," *Science*, vol. 330, no. 6007, pp. 1099–1102, 2010.

[98] M. Scott and T. Hwa, "Bacterial growth laws and their applications," *Current opinion in biotechnology*, vol. 22, no. 4, pp. 559–565, 2011.

[99] A. Y. Weisse, D. A. Oyarzun, V. Danos, and P. S. Swain, "Mechanistic links between cellular trade-offs, gene expression, and growth," *Proceedings of the National Academy of Sciences*, p. 201416533, 2015.

[100] R. Bürger, "On the maintenance of genetic variation: global analysis of kimura's continuum-of-alleles model," *Journal of mathematical biology*, vol. 24, no. 3, pp. 341–351, 1986.

[101] M. Doebeli and I. Ispolatov, "Complexity and diversity," *Science*, vol. 328, no. 5977, pp. 494–497, 2010.

[102] M. Doebeli and I. Ispolatov, "Chaos and unpredictability in evolution," *Evolution*, vol. 68, no. 5, pp. 1365–1373, 2014.

[103] C. C. Spencer, M. Bertrand, M. Travisano, and M. Doebeli, "Adaptive diversification in genes that regulate resource use in escherichia coli," *PLoS genetics*, vol. 3, no. 1, p. e15, 2007.

[104] A. Wagner, "Neutralism and selectionism: a network-based reconciliation," *Nature Reviews Genetics*, vol. 9, no. 12, pp. 965–974, 2008.

[105] B. Lehner, "Genes confer similar robustness to environmental, stochastic, and genetic perturbations in yeast," *PLoS One*, vol. 5, no. 2, p. e9035, 2010.

[106] S. F. Levy and M. L. Siegal, "Network hubs buffer environmental variation in saccharomyces cerevisiae," *PLoS biology*, vol. 6, no. 11, p. e264, 2008.

[107] J. R. Karr, J. C. Sanghvi, D. N. Macklin, M. V. Gutschow, J. M. Jacobs, B. Bolival, N. Assad-Garcia, J. I. Glass, and M. W. Covert, "A whole-cell computational model predicts phenotype from genotype," *Cell*, vol. 150, no. 2, pp. 389–401, 2012.

[108] H. Caswell, *Matrix population models*. Wiley Online Library, 2001.

[109] D. Molenaar, R. van Berlo, D. de Ridder, and B. Teusink, "Shifts in growth strategies reflect tradeoffs in cellular economics," *Molecular systems biology*, vol. 5, no. 1, p. 323, 2009.

[110] L. A. Segel and M. Slemrod, "The quasi-steady-state assumption: a case study in perturbation," *SIAM review*, vol. 31, no. 3, pp. 446–477, 1989.

[111] D. Herbert, R. Elsworth, and R. Telling, "The continuous culture of bacteria; a theoretical and experimental study," *Journal of General Microbiology*, vol. 14, no. 3, pp. 601–622, 1956.

[112] S. Friedberg, A. Insel, and L. Spence, "Linear algebra. 1997."

[113] U. Dieckmann and R. Law, "The dynamical theory of coevolution: a derivation from stochastic ecological processes," *Journal of mathematical biology*, vol. 34, no. 5-6, pp. 579–612, 1996.

[114] C. S. Holling, "The functional response of predators to prey density and its role in mimicry and population regulation," *Memoirs of the Entomological Society of Canada*, vol. 97, no. S45, pp. 5–60, 1965.

[115] P. Waltman, S. P. Hubbell, and S. B. Hsu, "Theoretical and experimental investigations of microbial competition in continuous culture," *Modeling and differential equations in biology. New York: Marcel Dekker*, pp. 107–192, 1980.

[116] S. H. Strogatz, *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*. Westview press, 2014.

[117] J. D. Orth, I. Thiele, and B. Ø. Palsson, "What is flux balance analysis?," *Nature biotechnology*, vol. 28, no. 3, pp. 245–248, 2010.

[118] J. S. Edwards, M. Covert, and B. Palsson, "Metabolic modelling of microbes: the flux-balance approach," *Environmental microbiology*, vol. 4, no. 3, pp. 133–140, 2002.

[119] F. J. Poelwijk, P. D. Heyning, M. G. de Vos, D. J. Kiviet, and S. J. Tans, "Optimality and evolution of transcriptionally regulated gene expression," *BMC systems biology*, vol. 5, no. 1, p. 128, 2011.

[120] H. H. McAdams, B. Srinivasan, and A. P. Arkin, "The evolution of genetic regulatory systems in bacteria," *Nature Reviews Genetics*, vol. 5, no. 3, pp. 169–178, 2004.

[121] A. S. Seshasayee, G. M. Fraser, M. M. Babu, and N. M. Luscombe, "Principles of transcriptional regulation and evolution of the metabolic system in e. coli," *Genome research*, vol. 19, no. 1, pp. 79–91, 2009.

[122] M. Heinemann and U. Sauer, "Systems biology of microbial metabolism," *Current opinion in microbiology*, vol. 13, no. 3, pp. 337–343, 2010.

[123] Wolfram Research, Inc., *Mathematica, 10.0*.

[124] L. F. Shampine, *Numerical solution of ordinary differential equations*, vol. 4. CRC Press, 1994.

[125] G. D. Battista, P. Eades, R. Tamassia, and I. Tollis, "Graph drawing: algorithms for the visualization of graphs," *Alan Apt*, 1999.

[126] P. Brémaud, *Markov chains: Gibbs fields, Monte Carlo simulation, and queues*, vol. 31. Springer Science & Business Media, 2013.

[127] M. Panju, "Iterative methods for computing eigenvalues and eigenvectors," *arXiv preprint arXiv:1105.1185*, 2011.

[128] W. H. Press, *Numerical recipes 3rd edition: The art of scientific computing*. Cambridge university press, 2007.

[129] L. J. Allen, *An introduction to stochastic processes with applications to biology*. CRC Press, 2010.

[130] J. A. Draghi, T. L. Parsons, G. P. Wagner, and J. B. Plotkin, "Mutational robustness can facilitate adaptation," *Nature*, vol. 463, no. 7279, pp. 353–355, 2010.

[131] P. F. Stadler and W. Schnabl, "The landscape of the traveling salesman problem," *Physics Letters A*, vol. 161, no. 4, pp. 337–344, 1992.

[132] P. Schuster, W. Fontana, P. F. Stadler, and I. L. Hofacker, "From sequences to shapes and back: a case study in rna secondary structures," *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 255, no. 1344, pp. 279–284, 1994.

[133] E. Van Nimwegen, J. P. Crutchfield, and M. Huynen, "Neutral evolution of mutational robustness," *Proceedings of the National Academy of Sciences*, vol. 96, no. 17, pp. 9716–9720, 1999.

[134] J. A. León and D. B. Tumpson, "Competition between two species for two complementary or substitutable resources," *Journal of theoretical Biology*, vol. 50, no. 1, pp. 185–201, 1975.

[135] B. Li and H. Smith, "How many species can two essential resources support?," *SIAM Journal on Applied Mathematics*, vol. 62, no. 1, pp. 336–366, 2001.

[136] M. M. Ballyk and G. S. Wolkowicz, "Exploitative competition in the chemostat for two perfectly substitutable resources," *Mathematical biosciences*, vol. 118, no. 2, pp. 127–180, 1993.

[137] M. M. Ballyk, C. C. McCluskey, and G. S. Wolkowicz, "Global analysis of competition for perfectly substitutable resources with linear response," *Journal of mathematical biology*, vol. 51, no. 4, pp. 458–490, 2005.

[138] M. Newman, *Networks: an introduction*. OUP Oxford, 2010.

[139] M. Hoffman, S. Suetens, U. Gneezy, and M. A. Nowak, "An experimental investigation of evolutionary dynamics in the rock-paper-scissors game," *Scientific reports*, vol. 5, 2015.

[140] Z. D. Blount, C. Z. Borland, and R. E. Lenski, "Historical contingency and the evolution of a key innovation in an experimental population of escherichia coli," *Proceedings of the National Academy of Sciences*, vol. 105, no. 23, pp. 7899–7906, 2008.

[141] J. R. Meyer, D. T. Dobias, J. S. Weitz, J. E. Barrick, R. T. Quick, and R. E. Lenski, "Repeatability and contingency in the evolution of a key innovation in phage lambda," *Science*, vol. 335, no. 6067, pp. 428–432, 2012.

[142] A. H. Melnyk and R. Kassen, "Adaptive landscapes in evolving populations of pseudomonas fluorescens," *Evolution*, vol. 65, no. 11, pp. 3048–3059, 2011.

[143] M. Travisano, J. A. Mongold, A. F. Bennett, and R. E. Lenski, "Experimental tests of the roles of adaptation, chance, and history in evolution," *Science*, vol. 267, no. 5194, pp. 87–90, 1995.

[144] K. N. Laland, F. J. Odling-Smee, and M. W. Feldman, "Evolutionary consequences of niche construction and their implications for ecology," *Proceedings of the National Academy of Sciences*, vol. 96, no. 18, pp. 10242–10247, 1999.

[145] N. Q. Balaban, J. Merrin, R. Chait, L. Kowalik, and S. Leibler, "Bacterial persistence as a phenotypic switch," *Science*, vol. 305, no. 5690, pp. 1622–1625, 2004.

[146] W. C. Ratcliff and R. F. Denison, "Individual-level bet hedging in the bacterium sinorhizobium meliloti," *Current Biology*, vol. 20, no. 19, pp. 1740–1744, 2010.

[147] M. M. Crane, I. B. Clark, E. Bakker, S. Smith, and P. S. Swain, "A microfluidic system for studying ageing and dynamic single-cell responses in budding yeast," 2014.

[148] D. F. Toupo and S. H. Strogatz, "Nonlinear dynamics of the rock-paper-scissors game with mutations," *Physical Review E*, vol. 91, no. 5, p. 052907, 2015.

[149] B. Sinervo and C. M. Lively, "The rock-paper-scissors game and the evolution of alternative male strategies," *Nature*, vol. 380, no. 6571, pp. 240–243, 1996.

[150] A. W. Miller, C. Befort, E. O. Kerr, and M. J. Dunham, "Design and use of multiplexed chemostat arrays," *JoVE (Journal of Visualized Experiments)*, no. 72, pp. e50262–e50262, 2013.

[151] K. Nishimura, T. Fukagawa, H. Takisawa, T. Kakimoto, and M. Kanemaki, "An auxin-based degron system for the rapid depletion of proteins in nonplant cells," *Nature methods*, vol. 6, no. 12, pp. 917–922, 2009.

[152] H. Alper, C. Fischer, E. Nevoigt, and G. Stephanopoulos, "Tuning genetic control through promoter engineering," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 36, pp. 12678–12683, 2005.

[153] A. M. Smith, T. Durbic, J. Oh, M. Urbanus, M. Proctor, L. E. Heisler, G. Giaever, and C. Nislow, "Competitive genomic screens of barcoded yeast libraries," *JoVE (Journal of Visualized Experiments)*, no. 54, pp. e2864–e2864, 2011.

[154] D. Button, "Nutrient uptake by microorganisms according to kinetic parameters from theory as related to cytoarchitecture," *Microbiology and Molecular Biology Reviews*, vol. 62, no. 3, pp. 636–645, 1998.

[155] S. Suzuki, T. Horinouchi, and C. Furusawa, "Prediction of antibiotic resistance by gene expression profiles," *Nature communications*, vol. 5, 2014.

[156] Y. Cao, D. T. Gillespie, and L. R. Petzold, "Efficient step size selection for the tau-leaping simulation method," *The Journal of chemical physics*, vol. 124, no. 4, p. 044109, 2006.

[157] K. K. Steffen, B. K. Kennedy, and M. Kaeberlein, "Measuring replicative life span in the budding yeast," *Journal of visualized experiments: JoVE*, no. 28, 2009.