

In Which The Fixation Probability Of A Superstar Is Determined

And A Contradiction In The Literature Is Addressed

by

Alastair David Jamieson-Lane

B.Sc., The University of Canterbury, 2012

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

The Faculty of Graduate and Postdoctoral Studies

(Mathematics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

October 2014

© Alastair David Jamieson-Lane 2014

Abstract

Population structures can be crucial determinants of evolutionary processes. For the spatial Moran process certain structures suppress selective pressure, while others amplify it (Lieberman *et al.* 2005 *Nature* **433** 312-316). Evolutionary amplifiers suppress random drift and enhance selection. Recently, some results for the most powerful known evolutionary amplifier, the superstar, have been invalidated by a counter example (Díaz *et al.* 2013 *Proc. R. Soc. A* **469** 20130193). Here we correct the original proof and derive improved upper and lower bounds, which indicate that the fixation probability remains close to $1 - 1/r^4 H$ for population size $N \rightarrow \infty$ and structural parameter $H > 1$. This correction resolves the differences between the two aforementioned papers. We also confirm that in the limit $N, H \rightarrow \infty$ superstars remain capable of providing arbitrarily strong selective advantages to any beneficial mutation, eliminating random drift. In addition, we investigate the robustness of amplification in superstars, and find that it appears to be a fragile phenomenon with respect to changes in the selection or mutation processes.

Preface

This project was initiated by Professor Christoph Hauert in response to the preprint of Díaz et al. [2013]. All mathematical investigation and proofs contained were constructed by Alastair Jamieson-Lane, with guidance provided by C. Hauert.

The paper upon which this thesis was based was originally drafted by Alastair J.L. Feedback was provided by C. Hauert.

Table of contents

Abstract	ii
Preface	iii
Table of contents	iv
List of tables	vi
List of figures	vii
Acknowledgements	viii
Dedication	ix
1 In which concepts are introduced	1
1.1 Context	1
1.2 Introduction to the Moran process on graphs	3
1.2.1 Superstars	6
2 In which prior work is discussed	9
2.1 Original proof: Lieberman <i>et al</i> (2005)	9
2.2 Counter example: Díaz <i>et al</i> (2013)	11
2.3 Initial comparison	13
3 In which a proof is described	17
3.1 Timescales	18
3.2 Top of stem	19
3.3 Middle of stem	19
3.4 Bottom of stem	23
3.5 Slow dynamics in reservoirs	23
3.6 Upper bound on fixation probability	23
3.7 Lower bound on fixation probability	24
3.8 Narrow window	25

Table of contents

4	In which error bounds are found	26
4.1	Initial conditions	26
4.2	Details of expected train length T .	27
4.2.1	Bounding T	27
4.2.2	Train collisions	29
4.3	Interaction of trains with root node	30
4.3.1	An upper bound on train success probability	32
4.3.2	A lower bound on train success probability	33
4.4	Bounds on fixation probabilities	35
4.4.1	Upper bound	35
4.4.2	Lower bound	36
4.5	Bringing it all together	39
5	In which some loose ends are tied up and others are left open	41
5.1	Selection & sequence	41
5.2	Mutations	43
5.3	Deleterious mutations, $r < 1$	43
6	In which concluding remarks occur	46
	Bibliography	48
 Appendices		
A	Programs	51
A.1	superstarSolve	51
A.2	SimpleChainTest	54

List of tables

2.1	Predictions based on Lieberman <i>et al</i> 's density argument, compared with the recorded density...	14
-----	---	----

List of figures

1.1	Spatial Moran process	4
1.2	The superstar...	6
2.1	An illustration of Díaz <i>et al</i> 's Markov chain argument.	12
2.2	Comparing simulation to density prediction..	15
3.1	Two possible histories of a train...	20
3.2	Grid showing collection of possible train states	21
4.1	Possible states of the end of the stem	31

Acknowledgements

I would like to thank my supervisor, Professor Christoph Hauert, for providing me with puzzles to be solved, and for guidance, both in mathematics, and in the more extended maze of academia.

I would also like to thank Dr. David Brydges for instruction in probability, and for providing a sounding board for my thoughts when I felt stuck. Thanks is also due to two anonymous reviewers who caught a flaw in the original proof, giving me the chance to correct it before publication.

In terms of funding, this work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Foundational Questions in Evolutionary Biology Fund (FQEB), grant RFP-12-10.

To the students of Mrs Lynn's class, grade three, division twelve, Lord Strathcona Elementary school, 2013-2014.

Good luck, have fun, don't grow up too quickly.

And never stop learning.

Chapter 1

In which concepts are introduced

1.1 Context

Populations evolve according to the principles of natural selection and random drift. The balance between the two competing processes is determined by numerous factors, including both population size and structure [Antal et al., 2006, Bürger and Lande, 1994, Nowak and May, 1992, Fu and Nowak, 2013]. The most malignant tumour is unlikely to cause harm if it arises in the outermost layer of skin and is easily brushed aside, and the most imperative model for climate change has limited influence until it has worked its way from a researchers desk, through the literature into policy making and public awareness. Position matters.

One of the simplest and most influential models of stochastic evolutionary processes in finite populations is the Moran process [Moran, 1962, Nowak, 2006]. It is based on an unstructured (or well-mixed) population of size N , where each individual is classified either as a resident (wild type) or mutant. Each type is assigned a constant fitness, which determines its propensity to reproduce. The fitness of wild types is normalized to 1 and mutants have fitness r . An advantageous mutant has $r > 1$, a disadvantageous mutant has $r < 1$ and a neutral mutant is indistinguishable in terms of fitness, $r = 1$. In every time step, an individual is randomly selected for reproduction with each individual's probability proportional to their fitness. The selected individual then produces a clonal offspring that replaces an individual, selected uniformly at random, in the population. This process is repeated until eventually the population has reached one of the homogenous states of all residents, if the mutant went extinct, or all mutants, if the mutant successfully took over the entire population [Moran, 1962, Nowak et al., 2004, Lieberman et al., 2005]. In both cases, the population has reached fixation, although for the sake of this document when talking of fixation we will be referring to fixation of mutants, unless stated otherwise. In the absence

of mutation, the two homogeneous states are absorbing.

The Moran process models evolutionary dynamics based on selection and random drift in finite populations: an advantageous mutant has a higher probability, but no guarantee, to reach fixation and, similarly, an inferior mutant is more likely to be eliminated, but not with certainty. Because of the simple (or rather, non-existent) topology of the original Moran process, all possible states of the system can be described by simply stating the current number of mutants. At any given time step, this number can either increase or decrease by one, with known probabilities. Any time step where the actual make up of the population changes must be associated with some interaction between a resident and a mutant – one replacing the other. Because mutants are always r times more likely to replace resident individuals than the converse, it can be shown that the chance of an increase in the number of mutant individuals is always r times greater than the chance of a decrease, leading to a “forward bias” of r regardless of the make-up of the population. These simple dynamics allow the system to be described as a random walk, where X_t denotes the current number of mutants, and the forward bias is always r . Simple martingale arguments can then be used to determine the probabilities of hitting either end of this random walk, from any starting point, and thus the fixation probabilities of either residents or mutants can be found for any given initial configuration. Of particular interest is the fixation probability, ρ_M , of a single mutant that arises in an otherwise homogeneous population of resident individuals:

$$\rho_M = \frac{1 - \frac{1}{r}}{1 - \frac{1}{r^N}}. \quad (1.1)$$

In the neutral limit, $r \rightarrow 1$, all individuals in the population are equally likely to end up as the single common ancestor, leading to a fixation probability of $1/N$.

The original Moran process ignores population structures – but this is easily addressed by arranging individuals of a population on a graph, such that each node refers to one individual and the links to other nodes define its neighbourhood. Maruyama [1970] and Slatkin [1981] conjectured that the fixation probability of a mutant in this Moran process on graphs remains unaffected by population structures. Lieberman et al. [2005] proved that this is indeed true for a broad class of structures and, in particular, holds for simple structures such as lattices or regular networks. At the same time, this classification indicated that fixation probabilities, ρ , may differ for some structures by tilting the balance between selection and random

drift. Evolutionary suppressors enhance random drift and suppress selection ($\rho < \rho_M$ for $r > 1$ and $\rho > \rho_M$ for $r < 1$), whereas evolutionary amplifiers enhance selection and suppress random drift ($\rho > \rho_M$ for $r > 1$ and $\rho < \rho_M$ for $r < 1$).

In recent years, various aspects of the Moran processes on graphs have been explored, including effects of population structures on fixation probabilities [Antal et al., 2006, Broom and Rychtář, 2008, Burton Voorhees, 2013], or fixation times [Joshua L. Payne, 2009, Frean et al., 2013], as well as computational techniques [Shakarjian and Roos, 2011, Fu et al., 2009]. However, the most surprising result remains that even perfect evolutionary amplification appears to be possible: “The superstar... [has] the amazing property that for large [population sizes] N , the fixation probability of an advantageous mutant converges to one, while the fixation probability of a disadvantageous mutant converges to zero.” [Lieberman et al., 2005].

A more recent paper Díaz et al. [2013] provided a counter example that contradicted the fixation probabilities reported in Lieberman et al. [2005]. In this thesis, I give a brief overview of the arguments in both papers, compare their predictions to some numerical results, and go on to identify the problem in the original proof and correct it. This yields new upper and lower bounds on the fixation probability for superstars. It is found that for any $r > 1$, a graph can be constructed such that ρ is arbitrarily close to 1, thus confirming the existence of perfect evolutionary amplification.

1.2 Introduction to the Moran process on graphs

Population *structure* can be represented by assigning *individuals* to nodes on a graph, with links representing each *individual's* neighbourhood¹. The Moran process on graphs follows the same procedure as the original Moran process except for the crucial difference that the offspring does not replace a random member of the entire population but rather replaces a *neighbour* of the reproducing individual, selected uniformly at random (Fig. 1.1).

On directed graphs, the offspring replaces a downstream neighbour by selecting one outgoing link uniformly at random. As before, the population has reached fixation once either one of the absorbing homogeneous states is reached. For any number of mutants, m ($0 < m < N$), the fixation prob-

¹Models also exist where nodes represent “islands” or some other unit of geographically isolated subpopulation. While interesting and certainly biologically relevant, this was the assumption of neither of the articles that this work is based on, thus is *not* what we will be assuming here.

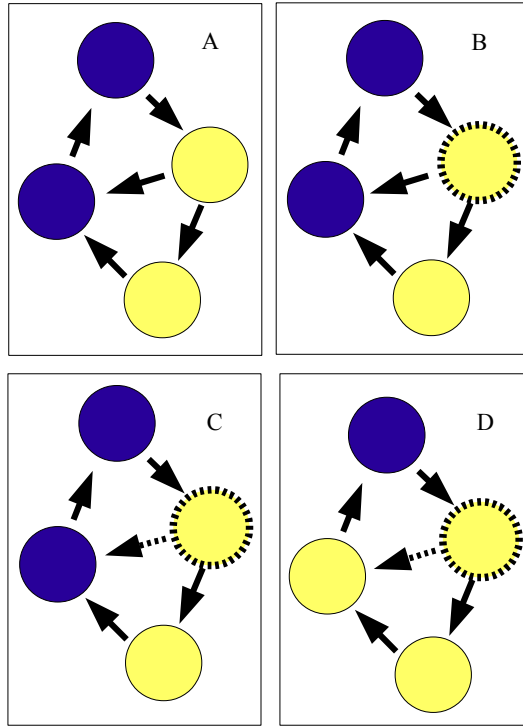


Figure 1.1: Spatial Moran process. (A) graph structure and initial distribution of residents (dark blue) and mutants (pale yellow). (B) selection: an individual (dashed outline) is selected to reproduce with a probability proportional to its fitness. (C) replacement: a downstream neighbour (dashed arrow) is randomly selected for replacement. (D) reproduction: the neighbour is replaced by the clonal offspring of the upstream reproducing individual.

abilities of residents and mutants are both non-zero on strongly connected graphs, i.e. graphs where every node can be reached from any other node through a series of moves between nodes that are connected by links (for directed graphs, only moves in the direction of the link are permitted). If a graph is not strongly connected, then the structure may prevent the spreading or elimination of a mutant type regardless of its fitness and hence the fixation probability for either or both types may be zero.

For the Moran process on graphs, the fixation probabilities are the same as in unstructured populations, c.f. Eq. (1.1), provided that the graph is a circulation [Lieberman et al., 2005]. A circulation is defined as a graph such

that the sum of weights of all outgoing links is equal to the sum of weights of all incoming links for each node. This means that every node has the same impact on the environment as the environment has on the node.

A graph is an evolutionary suppressor if the fixation probability of an advantageous mutant is less than for the original Moran process, $\rho < \rho_M$. The simplest example is a linear chain: a graph with a single source node, which connects to one end of a (directed) chain of nodes [Nowak et al., 2003]. Any mutation that does not occur at the top of the chain has no chance of reaching fixation. However, if the mutation occurs at the top node it will eventually reach fixation with certainty. Assuming that mutations arise spontaneously and are equally likely in any location, the resulting fixation probability is simply $1/N$, regardless of the mutant's fitness r . The linear chain is an example of a graph that is not strongly connected, because the source node cannot be reached from any node in the chain. Evolutionary suppressors are sometimes found when high fidelity copying is of paramount importance, such as in slowing down the somatic evolution of cancer [Nowak et al., 2003, Michor et al., 2004].

In contrast, an evolutionary amplifier is a graph which increases the fixation probability of advantageous mutants as compared to the original Moran process, $\rho > \rho_M$. The simplest evolutionary amplifier is the star graph: a single root node is connected to a reservoir of peripheral leaf nodes through bi-directional links. The fixation probability of a single mutant for $N \gg 1$ is [Lieberman et al., 2005, Broom and Rychtář, 2008]

$$\rho_0 \approx \frac{1 - \frac{1}{r^2}}{1 - \frac{1}{r^{2N}}}. \quad (1.2)$$

On the star, a mutant with fitness r has roughly the same fixation probability as a mutant with fitness r^2 would in an unstructured population. Thus, the fixation probability of beneficial mutations ($r > 1$) is enhanced, but for deleterious mutants ($r < 1$) it is reduced. Note that the fixation probability depends on where the single mutant arises. If the mutant is located in the root node then, for $N \gg 1$, it is almost certainly replaced in the next time step because one of the $N - 1$ reservoir nodes is selected for reproduction. However, if mutants arise at random, then for $N \gg 1$ they almost surely arise in the reservoir and the fixation probability is as specified in Eq. (1.2). Evolutionary amplifiers would seem to provide promising structures for tasks where strong selection is advantageous, such as in the adaptive immune system or in collaboration networks.

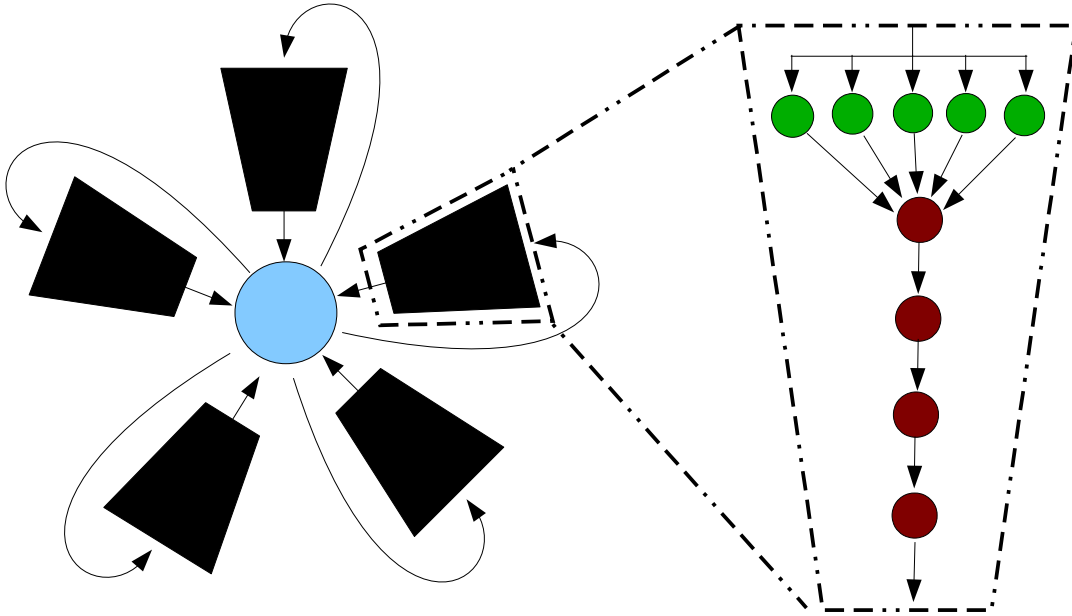


Figure 1.2: The superstar consists of three distinct types of nodes: the root node (pale blue), the reservoir nodes (green) and the stem nodes (dark red). The reservoir nodes connect to the start of the stem, the end of the stem connects to the root node and the root node connects to all reservoir nodes in each branch. The depicted superstar has $B = 5$ branches each with $L = 5$ reservoir nodes and a stem of length $H = 4$, which yields a total population size of $N = B(L + H) + 1 = 46$.

1.2.1 Superstars

The two most prominent features of the star graph are the large reservoir where changes occur on a slow time scale, and the bottleneck caused by the hub, where changes occur quickly. In particular, the bottleneck introduces a second level for selection to act upon. Lieberman et al. [2005] claim that this basic insight can be exploited to increase evolutionary amplification by elongating the bottleneck and providing further levels where selection can act. Superstars act as a more extreme version of the basic star, and are proposed as a way to increase evolutionary amplification further [Lieberman et al., 2005]. The superstar consists of a single root node surrounded by B branches (Fig. 1.2). Each branch consists of a large reservoir of L nodes feeding into one end of a directed chain of length H , the stem. The last

stem node in each branch feeds into the root node, which then connects to all reservoir nodes in every branch. The population size is thus given by $N = B(L + H) + 1$. Nodes are classified based on their locations on the graph. This classification is designed to simplify discussions but does not affect the rate of reproduction of the individual occupying the node. Lieberman et al. [2005] report the fixation probability for superstars with $L, B \gg H$ as

$$\rho_H \approx \frac{1 - \frac{1}{r^K}}{1 - \frac{1}{r^{KN}}}, \quad (1.3)$$

where $K = H + 2$ is a structural parameter and indicates the number of moves needed to reach any reservoir node from any other reservoir node. This is the number of levels selection can act upon. Consequently it is argued that a single mutant that arises in the reservoir of a superstar with fitness r has approximately the same fixation probability as a mutant with fitness r^K in an unstructured population. This result would then imply that by increasing the length of the stem, the fixation probability, ρ_H , of any advantageous mutant, $r > 1$, could be brought arbitrarily close to 1, indicating arbitrarily strong amplification or perfect selection.

Recently Díaz et al. [2013] provided a counter-example demonstrating that the fixation probability in Eq. (1.3) is too optimistic in the particular case of $H = 3$ and thus invalidated the proof in Lieberman et al. [2005]. In addition, they provided substantial simulation based evidence indicating that Eq. (1.3) also fails for higher values of H . For the counter-example they show that in the limit $N \rightarrow \infty$:

$$\rho_3 < 1 - \frac{1 + r}{2r^5 + r + 1}. \quad (1.4)$$

This upper bound reflects the probability that a mutant in a reservoir creates a second mutant in any reservoir before getting replaced by resident offspring. The fixation probability according to Eq. (1.3) grows faster with increasing r than Eq. (1.4) and for $r > 1.42$ results in a contradiction.

Upon arriving at the University of British Columbia I was asked to examine these conflicting results. In particular, I was tasked with:

- Reading both proofs in full detail.
- Determining whether the two results did in fact contradict each other

1.2. Introduction to the Moran process on graphs

- In the case of conflict, determining *precisely* where the conflict occurred.
- In the case of conflict, determine which of the two was correct
- Having determined which was correct, find the flaw in the incorrect proof

This initial work is covered in chapter 2. Having identified the critical flaw in the original proof [Lieberman et al., 2005], we then go on to consider a corrected proof, more accurate than [Lieberman et al., 2005], and more general than Díaz et al. [2013]. A general sketch of this proof is provided in chapter 3, with error terms and full rigor provided in chapter 4. Further explorations, and other loose ends are explored in chapter 5.

Chapter 2

In which prior work is discussed

The work in this thesis is concerned mainly with the work of two previous papers, namely Lieberman et al. [2005] and Díaz et al. [2013]. This chapter describes the basic arguments made in each paper, then goes on to discuss the significant conflicts between the two arguments, as well as some simulation results used to guide research near the start of the investigation. Where conflicts between our notation and that of the original authors arise, we use our own notation, so as to maintain internal consistency of this document.

2.1 Original proof: Lieberman *et al* (2005)

In this section, I describe the original argument, sketched in the supplementary material of Lieberman et al. [2005]. For the sake of completeness, I follow the proof as presented in Lieberman [2010, p. 47-54], which goes substantially beyond the original sketch.

The argument begins by first recognizing that, with overwhelming likelihood, the initial mutant (placed uniformly at random) will occur in the reservoir.

It is then argued that, if we have a series of populations in series, each downstream from the last, then, given one population with fixed mutant density d , the density of mutants in the following population will be $dr/(1 + d(r - 1))$. This can be viewed as the weighted fitness of mutants in the upstream population, dr relative to the total fitness of the upstream population $1 + d(r - 1)$. Density is not described precisely in the proof, but might be thought of as “the likelihood that a randomly chosen node, observed at a random time in the distant future, will contain a mutant”.

From this, induction is used to argue that, for population ν steps below our initial population, the density of mutants must be

$$d(\nu) = \frac{dr^\nu}{1 + d(r^\nu - 1)} \quad (2.1)$$

2.1. Original proof: Lieberman et al (2005)

This argument is extended to include the root node, thus leading to the root node having a mutant density of

$$d(K - 1) = \frac{dr^{K-1}}{1 + d(r^{K-1} - 1)} \quad (2.2)$$

Where $d = X_t/BL$ (X_t is the current number of reservoir mutants) and K is the length of the shortest cycle on the superstar (equivalent to $H + 2$)².

Lieberman et al. [2005] then gives a timing argument to argue that the density of mutants in the root has time to update such that it accurately reflects the density of mutants in the reservoirs (before any change occurs in the reservoirs). Unfortunately the timing argument as presented is flawed. The argument repeatedly uses the probability that a *mutant* will *eventually* pass a particular node in the stem before being replaced. In actual fact, the timespan needed is the *per time step* probability that a particular node is replaced by *any* upstream individual (since all upstream individuals can be considered as accurate measures of the upstream mutant density). As it happens, the corrected timing argument ends up showing that the time until update is faster than that predicted by Lieberman [2010], and thus this particular flaw in the original proof can be easily rectified without making any alterations to the surrounding arguments.

Next in the proof Eq. (2.2) is used to conclude that as long as the number of reservoir mutants is not too high, the probability that the number of reservoir mutants increases is very close to:

$$\frac{r}{N + X_t(r - 1)} \frac{dr^{K-1}}{1 + d(r^{K-1} - 1)} (1 - d). \quad (2.3)$$

While the probability of the number decreasing is close to

$$\frac{1}{N + X_t(r - 1)} \frac{1 - d}{1 + d(r^{K-1} - 1)} d. \quad (2.4)$$

The proof then considers a slower time scale, such that we only track time steps where the number of reservoir mutants changes (effectively conditioning on the occurrence of one of the above events). This allows them to

²For the benefit of any reader comparing this thesis to the original proof, we are forced to note that the original proof incorrectly refers to the root as node K . Because the reservoir is treated as level 0 it is believed that the root should be $K - 1$. Later in the the proof formulas are given treating the root as $K - 1$, leading to a slight inconsistency with previous working, but correcting for the previous mistake. Here we treat the root as node $K - 1$ throughout, in an attempt to follow the presumed intent of Lieberman et al. [2005] rather than the exact wording.

divide out all common terms in the above, normalising such that the probabilities sum to 1. The probability of reservoir mutants increasing rather than decreasing is found to be

$$\frac{r^K}{1+r^K}. \tag{2.5}$$

The total number of reservoir mutants is then treated as a random walk, with the above bias when X_t is small, and no bias otherwise. We omit the remainder of the proof, as it diverges substantially from anything we shall need to cover in this thesis.

2.2 Counter example: Díaz *et al* (2013)

Rather than attempting to form general results about all superstars, Díaz et al. [2013] instead focus on the particular case $K = 5$ (equivalently $H = 3$). They consider the behavior of the first mutant in the first branch, before any other reservoir mutants have been created. They show that at this early stage, the system has only 2^5 possible states, one for each combination of states of the initial mutant and its four downstream nodes (the last being the root). By the time any other node on the superstar is a mutant, we are no longer in the early stages of mutant propagation. Díaz et al. [2013] add one additional state to the above list, representing mutant propagation to another reservoir node. They then explicitly calculate the transition probabilities between all 33 states, resulting in a Markov chain with two absorbing states (extinction and propagation).

All transition probabilities are calculated *symbolically* in Díaz et al. [2013]. By excluding absorbing states, the problem is reduced to a system of 31 equations with 31 variables where each variable represents the probability of propagation occurring before extinction, given a particular starting state. This can be written in matrix notation as

$$M\rho = \rho + \mathbf{s}$$

where M is our transition matrix, ρ is a vector containing the eventual mutant propagation probabilities of each state, and \mathbf{s} is a vector containing the probability of our mutant strain propagating to a new branch *this time step*. This variable is referred to as “FS” in Díaz et al. [2013].

Mathematica is used to solve this system symbolically, resulting in an explicit expression for the probability that the mutant propagates to at least one other branch before extinction occurs. By taking the limit of large B

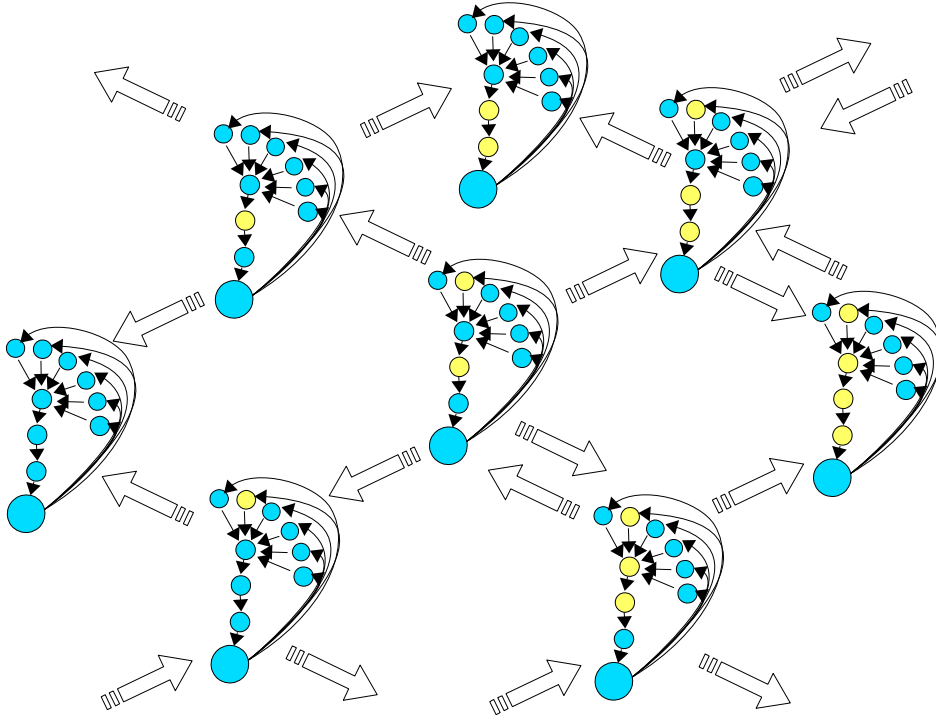


Figure 2.1: An illustration of Díaz et al. [2013]’s Markov chain argument. Possible states of the initial branch are listed, and all transition probabilities are calculated. Please note- many possible states and transitions are not represented, as doing so for all possible states would be impractical.

and L (M and L in Díaz et al. [2013]) all “dominated terms” are removed, resulting in

$$\rho_3 \leq 1 - \frac{1+r}{1+r+2r^5}, \quad (2.6)$$

an upper bound in direct conflict with the result stated in Lieberman et al. [2005] for sufficiently large r .

The paper then goes on to present results from a significant number of simulations, for a variety of r and K values. Generically, the results of the simulation disagree with the result stated in Lieberman et al. [2005], with a few minor exceptions for $r \approx 1$ and small K .

2.3 Initial comparison

Examining the proofs it can be seen that fundamentally the two proofs disagree on the probability that a single mutant will give rise to a second mutant. Any further arguments about fixation probability are made on the basis of this initial “mutant propagation probability”, and thus any attempt to resolve or explain this contradiction must begin with an explanation for their differing propagation probabilities.

My first step (after reading both proofs) was to implement the techniques of Díaz et al. [2013] numerically using MatLab. The code used can be found at the end of this thesis in appendix A.1. This was done primarily as a means of validating their program. If similar results were found using an independently written set of code, in a different programming language, using numerical rather than symbolic logic, this would strongly suggest that the Mathematica results were correct, and that any possibility for error would thus be limited to the conceptual part of the proof.

Running the code for $K = 5$, $B = 5000$, $L = 5000$ and $r = 1.5$, we find the probability of early extinction is 0.1415. We compare this to Eq. (2.6)’s prediction of 0.1413 and Eq. (2.5)’s prediction of 0.1164 .

It was observed that for any r value tested, the propagation probability found by the code converged to Díaz *et al*’s prediction for large L , B . This was considered strong evidence for the correctness of Díaz et al. [2013]’s code.

Next, I wished to test the validity of Lieberman et al. [2005]’s density argument. This I did by writing a simple java program to simulate the behavior *of a single branch* of my superstar. The small size of this subsystem allowed me to avoid many of the computational limitations encountered by simulators who attempted to analyze the entire system from $t = 1$ all the way to fixation. By only dealing with a single branch, the data structures required were also significantly simplified. This code can be found in appendix A.2.

Running this system for several days, for a variety of parameter values, I was able to extract the mutant density for all nodes in the stem, as well as the correlation between the states of neighbouring nodes. It was soon observed that the results of the simulation did not agree with the predictions made in Lieberman et al. [2005].

Table 2.1 contains some sample results.

2.3. Initial comparison

Node	Prediction $\times 10^{-5}$	Simulation $\times 10^{-5}$	Correlation
1	7.499	2.538	0.600316261771
2	11.24	3.045	0.719631574002
3	16.87	3.452	0.775873379784
4	25.30	3.802	0.807957544715
5	37.95	4.098	0.829692206727
6	56.92	4.342	0.844752636008
7	85.35	4.536	0.854818725014
8	127.9	4.683	0.862140794289
9	191.8	4.793	0.867951883031
10	287.4	4.879	0.871855826581
11	430.6	4.944	0.874703301922
12	644.5	4.997	0.877706319947
13	963.7	5.046	0.879744964568

Table 2.1: Predicted mutant density based on Lieberman *et al*'s density argument, compared with the recorded density from the ChainTest program in A.2. Parameter values $L = 20000$, $r = 1.5$, $K = 15$. The results are recorded at time $t = 67108864000$ in the special and condensed time frame of the program. Predictions are calculated using Eq. (2.1). When we refer to correlation here we mean “the probability that node $i + 1$ contains a mutant, given that i contains a mutant”

It turns out that Eq. (2.2) is calculated under the implicit assumption that the states of neighbouring nodes are independent of one another. It implies that whenever mutants are rare the probability of a mutant reproducing and replacing another mutant must also be low, even though the states of neighbouring nodes are highly correlated.

We might hope that taking neighbour-neighbour correlations into account will give us a more accurate induction formula. If we consider M_i , the state of node i being a mutant at a given time step, and R_i the state of node i being a resident type individual at a given time step, we can see that trivially:

$$\mathbb{P}(M_{i+1}) = \mathbb{P}(M_{i+1}|M_i)\mathbb{P}(M_i) + \mathbb{P}(M_{i+1}|R_i)\mathbb{P}(R_i)$$

because R and M are the only possible states. We might also remember that, in any given position, for every resident that gets replaced by a mutant, one mutant must be replaced by a resident (a lightbulb must be turned off before it can be turned back on again) . Because these two events must alternate

2.3. Initial comparison

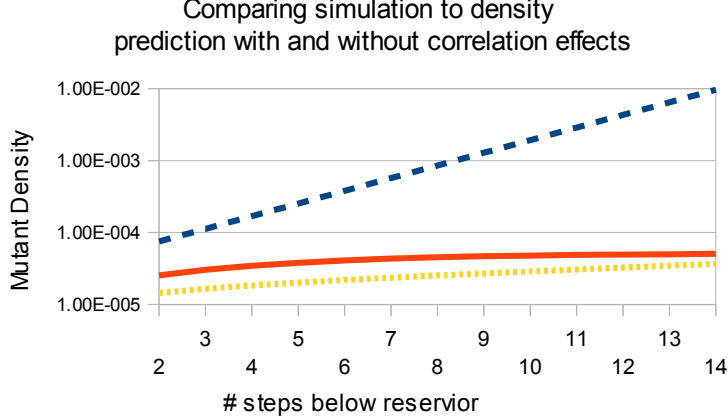


Figure 2.2: Here we compare the density recorded during our simulation (solid orange line) with density predictions made either accounting for (yellow, fine dash) or not accounting for (blue, long dash) the effects of neighbour-neighbour correlation. When accounting for correlation we use Eq. (2.7), and when ignoring it we use Eq. (2.1). It is observed that accounting for correlation in this most simple manner significantly improves the resulting density prediction, but still results in error.

with one another, and can get at most 1 out of sync, we can predict that in the long term the *probability* of these two events must equal. To calculate these probabilities we multiply the probability our graph is in a suitable state (mutant following resident, or resident following mutant) by the per time step probability that the upstream member of the pair reproduces. This gives:

$$\frac{r}{N}\mathbb{P}(R_{i+1}|M_i)\mathbb{P}(M_i) \approx \frac{1}{N}\mathbb{P}(M_{i+1}|R_i)\mathbb{P}(R_i)$$

Where here we use $1/N$ to approximate the fitness normalisation factor.

Taken together these give:

$$\mathbb{P}(M_{i+1}) \approx \mathbb{P}(M_i) [\mathbb{P}(M_{i+1}|M_i) + r(1 - \mathbb{P}(M_{i+1}|M_i))]. \quad (2.7)$$

This we can use to inductively predict the mutant density for nodes far down the chain (see Fig. 2.3), thus illustrating the importance of this correlation.

The “corrected” density formula identified gives a better prediction, but still differs significantly from our experimental observations – possibly due

2.3. Initial comparison

to failure to taken long distance correlation into account. Worse still, it can not be computed without using correlation data, which in the above example was collected *from the simulation itself*. Thus, its predictive power is largely post hoc. It serves as an illustration of the importance of correlation, but is of little use in *predicting* long term behavior. In order to truly predict the behavior of mutants in the stem, and hence calculate the overall fixation probability of superstar, we will need far more sophisticated arguments.

Chapter 3

In which a proof is described

Having determined that neighbour-neighbour correlation along the stem is an important feature of the system, it is clear this must be explicitly modeled if we wish to find the true fixation probability. Despite its rigor, the structure of Díaz et al. [2013] does not readily lend itself to extension to arbitrary H (although extension to any *particular* H is straight forward, but computationally expensive as H becomes large). Lieberman et al. [2005], while flawed in its details, is well structured, and was used as a starting point in what follows. Many of Lieberman *et al*'s steps are echoed - although the results and argument for each step often differ.

In this chapter we lay out a proof that in the asymptotic case the true fixation probability is

$$1 - \frac{1}{r^4(H-1)(1-1/r)^2} \leq 1 - \frac{1}{r^{4T}} \leq \rho_H \leq 1 - \frac{1}{1+r^{4T}} \leq 1 - \frac{1}{1+r^4H} \quad (3.1)$$

where T is “train length”- a value to be defined in section 3.3. In order to match previous work as closely as possible, it is assumed that there are many branches, B , and large reservoirs, L . Additionally, mutations are assumed to be beneficial, $r > 1$. As has been argued previously, if mutations arise spontaneously and with equal probability at any node, then the initial mutant almost certainly arises in a reservoir node, because reservoir nodes vastly outnumber nodes of all other types. As in both Díaz et al. [2013] and Lieberman et al. [2005], we study the dynamics of a single branch in detail, and use this to determine the much slower dynamics of changes in the reservoirs.

On occasion, it will prove useful to refer to the total fitness of all individuals in our superstar at a given time, we label this F_t . $N < F_t < rN$. All instances of F_t cancel throughout the proof and hence the exact value is not tracked.

This chapter sketches out the main thrust of the proof, following (roughly) the argument as it was created. For simplicity in sketching it is assumed $B = L \approx \sqrt{N}$. In the following chapter such simplifying assumptions are

dropped, and due attention is given to error bounds, taking limits simultaneously, and other omitted details.

3.1 Timescales

Different nodes get updated at different rates. More precisely, any given node of interest is updated if one of its upstream neighbours reproduces and the node of interest is chosen for replacement. Thus, individuals in nodes with high (weighted) in-degrees, tend to be short lived, while individuals in nodes with low (weighted) in-degrees, tend to be long lived. Here we follow the convention that each node has a total weighted out-degree of 1, and that weight is evenly distributed among outgoing links (each link has weight $1/k$ where k is the number of links).

Assuming $1 \approx r \ll N$, every node is selected for reproduction with probability $\approx 1/N$. The root node has an in-degree of B and all its upstream neighbours have out degrees of 1, hence it updates with a probability close to $B/N \approx 1/\sqrt{N}$. Similarly, reservoir nodes are replaced with probability of approximately $1/N^2$, the first stem node with probability on the order of $1/\sqrt{N}$, and all other stem nodes with probability of approximately $1/N$. For $N \gg 1$, this results in three different timescales: the slowest for reservoir nodes that get replaced, on average, only once in N^2 time steps; an intermediate timescale for the stem nodes (with the exception of the first node of each stem), which get replaced once in N time steps; and a fast timescale for the root node as well as the first stem node in each branch, which update once in \sqrt{N} time steps, respectively.

For $N \gg 1$, it is possible to separate the three timescales and analyze the dynamics of the different types of nodes individually. More specifically, this allows us to focus on the intermediate timescale associated with the dynamics in the stem, while treating the state of the slowly updating reservoir nodes as constant. Because the fast updating nodes are unlikely to reproduce twice without first being replaced many times themselves, we can treat them as random variables, with some probability of being a mutant on any given time step.

In the following, we derive the evolutionary dynamics for the top, middle and bottom of the stem in a single branch. The results determine the slow dynamics of reservoir nodes and describes the early stages of the invasion process, when mutants are rare among the reservoir nodes. This allows us to derive upper and lower bounds on the fixation probabilities.

3.2 Top of stem

The first node of the stem gets replaced on the fast time scale, which allows us to treat its state as a random variable. Out of all L upstream neighbours in the reservoir of the corresponding branch, only one is a mutant. Hence, at any given time step, the top node is occupied by a mutant with probability close to r/L . This mutant reproduces with a probability r/F_t and hence the probability that a mutant is placed in the second stem node is approximately $r^2/(F_t L)$ in each time step (for error terms, see 4.1). For reasons which will soon become apparent, this event is referred to as “train creation”.

3.3 Middle of stem

We refer to the collection of nodes that update on the medium time scale (that is, all stem nodes except the first) as the stem body. Simulations nicely illustrate the dynamics along the body of the stem: clusters of mutants begin at the top of the stem, then grow and move along the stem. In the following, we refer to these clusters as *trains*. A train moves forward and increases in length whenever the front mutant reproduces, which happens at a rate r/F_t . A train shrinks whenever a resident reproduces and replaces the back end of the train, which occurs at a rate $1/F_t$, see Fig. 3.1. Thus, as the train moves along the stem, the train length for beneficial mutants increases, on average.

Note that for small superstars with a single node in the stem body, which corresponds to $H = 2$ (or $K = 4$), the two stem nodes can be treated as effectively uncorrelated, due to their differing timescales. No train argument is needed. However, for $H > 2$ this assumption breaks down and results in an overestimation of the fixation probabilities. This explains why the predictions of Lieberman et al. [2005] hold in all cases they simulated, but break down for larger H .

In order to link the stem dynamics to the slow timescale of reservoir nodes, we need to know the expected train length, T , when the train first reaches the root end of the stem.

At any time, t , the state of a train can be described by two integers: A_t and Z_t . Here A_t refers to the position of the mutant at the front of the train, and Z_t refers to the position of the resident directly behind the train. The current length of the train is thus given by $A_t - Z_t$. Because in most time steps no change occurs in this particular stem, we consider the process in a condensed time frame which only accounts for events that change the

3.3. Middle of stem

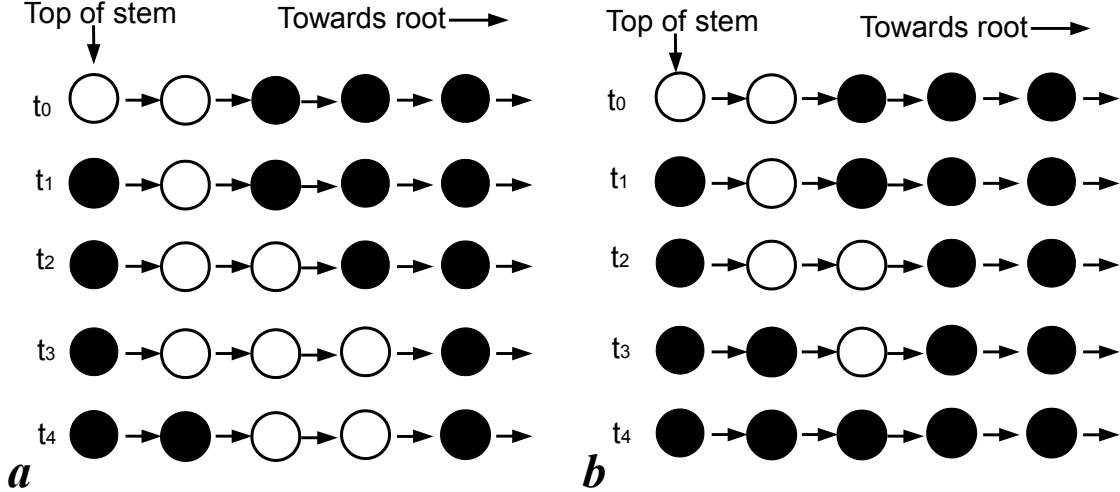


Figure 3.1: Two possible histories of a train of mutants (white) proceeding along a stem filled with residents (black). **a** We begin with two mutants (t_0). The top node is quickly replaced by a resident (on the fast time scale) (t_1). Some time later the remaining mutant reproduces (t_2), and then the new top node reproduces again (t_3). Finally we lose a single mutant from the back of our train (t_4). This general growing pattern applies whenever $r > 1$. **b** We begin with two mutants (t_0), and immediately lose the back mutant of our train (t_1). The front of the train reproduces, creating a second mutant (t_2), but both fall prey to bad fortune (or low fitness) and are removed (t_3, t_4). This behaviour is likely when $r < 1$, but even for beneficial mutations many trains do not reach the end of the stem.

state of the train. On this timescale A_t increases with probability $r/(1+r)$ while Z_t increases with probability $1/(1+r)$. Thus, for beneficial mutants the length tends to increase over time. If at any time $Z_t \geq A_t$ the train has vanished and the stem is cleared of mutants. In this case, we say that the train, which “arrives” at the end of the stem, has length zero.

In order to determine the expected train length, T , we consider the above process on a grid, where the horizontal axis represents the position of the front, A_t , and the vertical axis the back of the train, Z_t . Each point on the grid and below the diagonal, $A_t = Z_t$, represents a possible configuration of a train in the stem, see Fig. 3.2. All other points represent invalid configurations, which we refer to as ghost states. For each train, the initial

3.3. Middle of stem

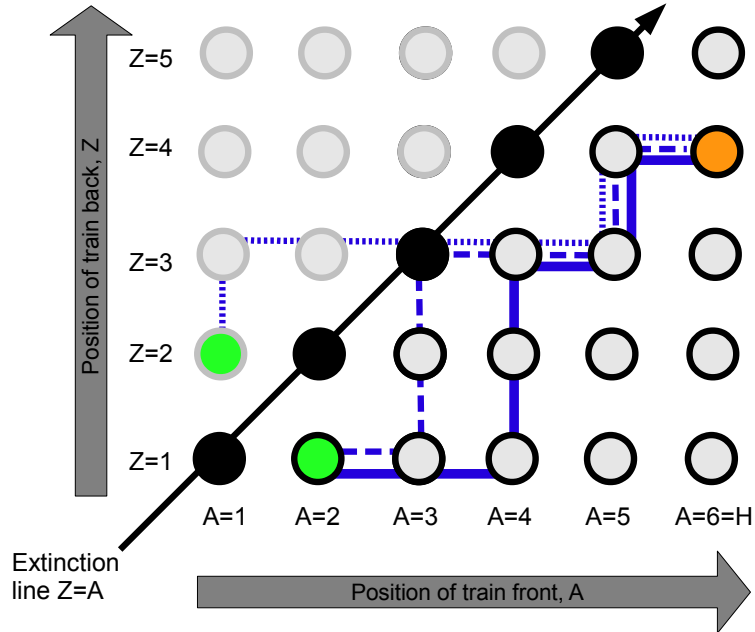


Figure 3.2: Grid showing collection of possible train states. Permitted states (black outline), ghost states (grey outline) and extinction states (black fill) as well as a number of possible paths from our initial state (black outline, green fill) to a sample end state (black outline orange fill). Depicted are a permitted path (continuous), an invalid path (leading to extinction, long dash) and the associated “ghost path” from the reflection of our initial state to the sample end state (fine dash).

configuration is $(A_0, Z_0) = (2, 1)$, that is, the second stem node is a mutant, while the state of the first stem node is a resident. Note that this assumes that the top of the stem is replaced by a resident on the fast timescale which leads to a slight underestimate.

Each train produces a path on the grid that originates in (A_0, Z_0) and ends when the train reaches its destination, $A_\tau = H$. Here τ refers to the train's time of arrival. If at any point in time $A_t \leq Z_t$ then this represents an *invalid* path because the train has gone extinct part way down the stem. The expected train length, T , is the weighted average over all paths, with invalid paths being considered as having length zero. The number of valid paths to any given end state (H, Z_τ) , can be calculated using the reflection principle

3.3. Middle of stem

[Koralov and Sinai, 2007, p. 88], which describes a one-to-one correspondence between *invalid* paths and “ghost paths” starting from (Z_0, A_0) . The trajectory of a ghost path is the reflection of the corresponding invalid path along the diagonal $A_t = Z_t$, up to the point where the invalid path touches $A_t = Z_t$ for the first time. From then on the ghost path and the remainder of the invalid path coincide, see Fig. 3.2. In order to calculate the expected train length, T , we consider all possible paths ending in $A_\tau = H$ subtract the number of *invalid* paths, to obtain the number of *valid* paths leading to a given end state. The number of ghost paths corresponds to all paths starting from (Z_0, A_0) , the reflection of (A_0, Z_0) and hence the name of the method. Having counted the number of valid paths reaching a particular end state we then weigh each path by its probability, multiply each end state by the corresponding train length, and sum over all paths to find:

$$T = (1 - \alpha)^{H-2} \sum_{z=1}^{H-1} (H - z) \alpha^{z-1} \left[\binom{H-3+z-1}{z-1} - \binom{H-3+z-1}{z-2} \right] \quad (3.2)$$

with $\alpha = 1/(1 + r)$. We assume beneficial mutations, $r > 1$, such that $0 < \alpha < 1/2$. All paths require $H - 2$ steps that increase A_t from the starting point at 2 to the end point at H , each of these steps occur with probability $1 - \alpha$. The combinatorial sum then accounts for every possible series of increments that Z_t might undergo along all valid paths. In particular, the index variable z indicates the final position of the tail of the train and hence $H - z$ specifies the train length. The tail starts at 1 and, for any valid path, reaches at most $H - 1$. Because we are interested in the length of the train at the moment of arrival, the final step must be an increment of A_t , this acts to limit the number of possible paths, lowering our binomials significantly. It follows that $z = H - 1$ has zero valid paths – a reassuring result as we know that no train could possibly have length one at the moment of its arrival. Note that we have used the convention that $\binom{n}{k} = 0$ for $k < 0$, which applies only if the tail remains at $Z_t = 1$ and admits only a single valid path.

For $H \geq 2, r > 1$ simple bounds for T exist (see section 4.2.1):

$$(H - 1) \left(1 - \frac{1}{r}\right)^2 \leq T \leq H. \quad (3.3)$$

This indicates that for $r > 1$ the expected train length T grows approximately linearly with increasing stem length H .

3.4 Bottom of stem

Whenever a train reaches the root end of the stem, its mutants compete with the resident nodes from the other branches to occupy the root node. Since the root node is updated on the fast timescale we can again treat its state as a random variable. Thus, once a train has reached the root end of the stem, the root node is a mutant with probability close to r/B . As long as the train sits at the root end of the stem, the probability in any given time step that the root node is a mutant and reproduces is close to $r^2/(F_t B)$. However, the train is simultaneously being eroded from behind, with train mutants being replaced with probability $1/F_t$. Thus, the train remains at the root end for TF_t time steps, on average. Put together, this means that any given train succeeds in producing a second mutant in any reservoir with a probability close to $r^2 T/B$ (for detailed error bounds see 4.3).

3.5 Slow dynamics in reservoirs

At any given time step, the probability of losing the initial mutant in the reservoir is close to $1/(F_t B L)$. In order to calculate the probability of creating a new reservoir mutant, we note that for each new reservoir mutant a successful train must have been launched, hence it suffices to find the per time step probability of launching a successful train. This can be easily calculated by simply multiplying the per time step probability of launching a train ($r^2/(F_t L)$, see section 3.2) by the probability that a given train is successful ($r^2 T/B$, see section 3.4), yielding approximately $r^4 T/(F_t B L)$.

Thus, the probability to eventually go from one to two mutants in the reservoirs, as opposed to losing the initial mutant, is close to

$$\frac{r^4 T}{1 + r^4 T}. \quad (3.4)$$

Since T can be made arbitrarily large (by increasing the stem length H , see Eq. (3.3)), the transition from one to two mutants becomes almost certain and, similarly, the probability of losing the initial mutant becomes vanishingly small.

3.6 Upper bound on fixation probability

To find an upper bound on the fixation probability, ρ_H , we note that before our mutant can reach fixation, the superstar must first transition from a

3.7. Lower bound on fixation probability

state with one mutant in a reservoir to a state with two mutants in the reservoirs (this is the basis of Díaz et al. [2013]). Thus, an upper bound on this transition probability serves as an upper bound on the mutant fixation probability for the system. Moreover, the upper bound can be made independent of T by assuming that all trains have the maximum possible train length. Thus, in the limit of large B and L we find

$$\rho_H \leq 1 - \frac{1}{1 + r^4 T} \leq 1 - \frac{1}{1 + r^4 H}. \quad (3.5)$$

For any given H, r we can find T explicitly using Eq. (3.2). In particular, we note that for $H = 3$, we find $T = 2r/(r + 1)$, thus recovering the upper bound identified in Díaz et al. [2013].

3.7 Lower bound on fixation probability

We find a lower bound on the fixation probability using the same approach as Lieberman [2010, p. 47-54], although the results we reach are very different. We approximate the dynamics of the system with a random walk, and then calculate the fixation probability on our random walk. This random walk has a forward bias given by Eq. (3.4) as long as mutants are rare, and we assume no forward bias otherwise. Because even for larger numbers of mutants the forward bias persists (but there is no simple way to quantify the bias) we obtain a lower bound of the fixation probability, ρ_H .

For any finite number of steps, a sufficiently strong initial bias would suffice to ensure that the random walk eventually reaches fixation with high probability. However, the limit $N \rightarrow \infty$ also requires an arbitrarily large number of forward steps. In order to resolve the interplay between these two limiting behaviours we set up a martingale and apply the optional stopping theorem [Klenke, 2006, p. 210] (see 4.4.2 for details).

In the limit of large B and L we find:

$$\rho_H \geq \hat{\rho}_H = 1 - \frac{1}{r^4 T} \geq 1 - \frac{1}{r^4 (H - 1) (1 - 1/r)^2}. \quad (3.6)$$

Once again we note that for any given H, r we can find T explicitly, Eq. (3.2).

3.8 Narrow window

Taken together, Eq. (3.6) and Eq. (3.5) give us narrow bounds on the possible values of fixation probability. For $B = L$ we find in the limit $B \rightarrow \infty$:

$$1 - \frac{1}{r^4(H-1)(1-r^{-1})^2} \leq 1 - \frac{1}{r^4T} \leq \rho_H \leq 1 - \frac{1}{1+r^4T} \leq 1 - \frac{1}{1+r^4H}. \quad (3.7)$$

Chapter 4

In which error bounds are found

Despite any intuitive appeal the previous chapter may have, and to some extent *because* of this, it is necessary to keep a most careful eye on any and all error terms that may arise. Previous dealings with superstars have demonstrated how delicate the balance is between the various limits we are dealing with. Our own investigations were occasionally led astray by the most innocuous of assumptions. Thus it is that we proceed with caution.

In this chapter, more formal arguments are given, to back up the intuitive notions in the previous chapter. We find exact error bounds, ϵ_i , on the numerous types of error that finite L and B introduce, and show that all such errors tend to zero. At the end of the chapter, all such error terms are collated to create the *finite case* upper and lower bounds for the fixation probability on superstars.

4.1 Initial conditions

If mutations arise spontaneously and with equal probability in any node then the initial mutant arises in a reservoir node with probability $BL/(BL + 1 + HB)$. This probability can be made arbitrarily close to one, for suitably large L . The mutant arises in a stem or root node with probability

$$\epsilon_0 = \frac{1 + HB}{BL + 1 + HB}. \quad (4.1)$$

Thus, the final fixation probability, ρ_H (see Eq. (6.1)), will need to include a $1 - \epsilon_0$ factor to account for this possibility. Because ϵ_0 is only used to derive the lower bound on ρ_H , assuming extinction of all mutants not arising in a reservoir node preserves inequalities. By doing this we are effectively ignoring the small possibility that an invading mutant placed in stem nodes or the root node could still reach fixation.

The first node of the stem gets replaced on the fast time scale, which allows us to treat its state as a random variable. However, at early stages

4.2. Details of expected train length T .

of invasion, only one of the L upstream neighbours is a mutant. Hence, at any given time step, the top node is occupied by a mutant with probability $r/(L - 1 + r) = (1 - \epsilon_1)r/L$, where

$$\epsilon_1 = \frac{r - 1}{L + r - 1}. \quad (4.2)$$

This mutant reproduces with a probability r/F_t and hence the probability that a mutant is placed in the second stem node is $r^2/(F_t(L + r - 1))$ in each time step. However, we need to account for the possibility that the initial mutant in the reservoir is replaced *before* the first node in that stem. On a given time step, the chance that the reservoir mutant is replaced by a resident is less than $1/(F_tBL)$. Conversely, the probability of the first node in the chain being replaced exceeds L/F_t . Thus the chance that the initial mutant is replaced before the first node in the chain is

$$\epsilon_2 < 1/(1 + BL^2). \quad (4.3)$$

For our proof we assume that the first node in any chain can be treated as a random variable. The above error terms ϵ_1 and ϵ_2 account for the slight simplifications we make in doing so.

4.2 Details of expected train length T .

In section 3.3, we showed that the expected length of a train in a stem was

$$T = (1 - \alpha)^{H-2} \sum_{z=1}^{H-1} (H - z)\alpha^{z-1} \left[\binom{H - 3 + z - 1}{z - 1} - \binom{H - 3 + z - 1}{z - 2} \right] \quad (4.4)$$

In this section, we identify upper and lower bounds on this expectation value, and deal with another source of minor error.

4.2.1 Bounding T

Finding an upper bound on T is very straight forward: because of the structure of the graph $H > T$.

To find a useful lower bound we are forced to resort to algebraic manipulation. Various binomial coefficient identities are used throughout. It is assumed that $r > 1$ and hence $0 < \alpha < 1/2$

$$T = (1 - \alpha)^{H-2} \sum_{z=1}^{H-1} (H - z)\alpha^{z-1} \left[\binom{H + z - 4}{z - 1} - \binom{H + z - 4}{z - 2} \right]$$

4.2. Details of expected train length T .

using Pascal's rule

$$= (1 - \alpha)^{H-2} \sum_{z=1}^{H-1} (H - z) \alpha^{z-1} \left[\binom{H+z-3}{z-1} - 2 \binom{H+z-4}{z-2} \right]$$

splitting sum

$$= (1 - \alpha)^{H-2} \sum_{z=1}^{H-1} (H - z) \alpha^{z-1} \binom{H+z-3}{z-1} - 2\alpha(1 - \alpha)^{H-2} \sum_{z=1}^{H-1} (H - z) \alpha^{z-2} \binom{H+z-4}{z-2}$$

changing second summation to obtain lower bound

$$\geq (1 - \alpha)^{H-2} \sum_{z=1}^{H-1} (H - z) \alpha^{z-1} \binom{H+z-3}{z-1} - 2\alpha(1 - \alpha)^{H-2} \sum_{z=2}^H (H + 1 - z) \alpha^{z-2} \binom{H+z-4}{z-2}$$

merging sums and relabeling indices

$$= (1 - 2\alpha)(1 - \alpha)^{H-2} \sum_{z=0}^{H-2} (H - 1 - z) \alpha^z \binom{H+z-2}{z}$$

expanding factor

$$= (1 - 2\alpha)(1 - \alpha)^{H-2} \sum_{z=0}^{H-2} (2(H - 1) - (H - 1) - z) \alpha^z \binom{H+z-2}{z}$$

using the combinatorial identity $(n+k) \binom{n+k-1}{k} = (n+k) \binom{n+k-1}{n-1} = n \binom{n+k}{n} = n \binom{n+k}{k}$

$$= (H - 1)(1 - 2\alpha)(1 - \alpha)^{H-2} \sum_{z=0}^{H-2} \alpha^z \left[2 \binom{H+z-2}{z} - \binom{H+z-1}{z} \right]$$

extending the sum to ∞ can only decrease the lower bound because $2 \binom{n+k}{k} - \binom{n+k+1}{k} = \binom{n+k}{k} - \binom{n+k}{k-1} \leq 0$ for $k > n$ and $(1 - 2\alpha) > 0$

$$\geq (H - 1)(1 - 2\alpha)(1 - \alpha)^{H-2} \sum_{z=0}^{\infty} \alpha^z \left[2 \binom{H+z-2}{z} - \binom{H+z-1}{z} \right]$$

4.2. Details of expected train length T .

using $(1 - \alpha)^{-n-1} = \sum_{k=0}^{\infty} \alpha^k \binom{n+k}{k}$ since $|\alpha| < 1$

$$= (H - 1)(1 - 2\alpha)(1 - \alpha)^{H-2} [2(1 - \alpha)^{1-H} - (1 - \alpha)^{-H}] = (H - 1) \left(\frac{1 - 2\alpha}{1 - \alpha} \right)^2.$$

And so we have

$$T \geq (H - 1) \left(1 - \frac{1}{r} \right)^2. \quad (4.5)$$

Hence, for $r > 1$, the expected train length, T , can be made arbitrarily long by choosing a suitably long chain, H .

4.2.2 Train collisions

The above derivation of the expected train length neglects the possibility that two trains may collide and merge, which introduces a source of error. Any train collision effectively reduces the train length of the first train because it is eradicated from behind at a greater rate than expected. Therefore, collisions *decrease* the expected train length T – despite the fact that merging trains may lead to longer overall lengths – hence Eq. (3.2) overestimates the expected train length. A lower bound for T is obtained by assuming that the second train completely eradicates the first train. In order to find this lower bound we need to determine the probability that a train collision occurs. An upper bound on this collision probability is given by the probability that a second train is generated while another train is still occupying the stem and can be formulated in terms of a negative binomial distribution where the generation of a new train counts as a “success” while a decrease in length of the existing train in the stem counts as a “failure”. If the rear of the train has incremented H times before a new train is generated, then the trains can not even co-habit the stem, let alone collide.

In each time step a new train is generated with probability $r^2/(F_t(L + r - 1))$ whereas the probability that the existing train length decreases, i.e. the resident directly behind the train reproduces, is at least $1/F_t$ (exactly $1/F_t$ along the stem, but greater for the first stem node). After H failure events we know that the stem must be cleared and contain only residents. Therefore, train collisions occur at most with the probability that a new train is generated prior to H failure events:

$$P(\text{no } 2^{\text{nd}} \text{ train} \geq \left(1 - \frac{r^2}{L + r - 1 + r^2} \right)^H > 1 - H \frac{r^2}{L + r - 1 + r^2}. \quad (4.6)$$

4.3. Interaction of trains with root node

The inequality in Eq. (4.6) results from expanding and truncating the power term. This yields an alternating sum and, for sufficiently large L , the absolute value of the terms is strictly decreasing. The chance that a second train is launched while another one still occupies the stem is at most

$$\epsilon_3 = \frac{Hr^2}{L + r - 1 + r^2} \quad (4.7)$$

and becomes small for $L \gg Hr^2$. Thus, the true expected train length lies between $T(1 - \epsilon_3)$ and T .

4.3 Interaction of trains with root node

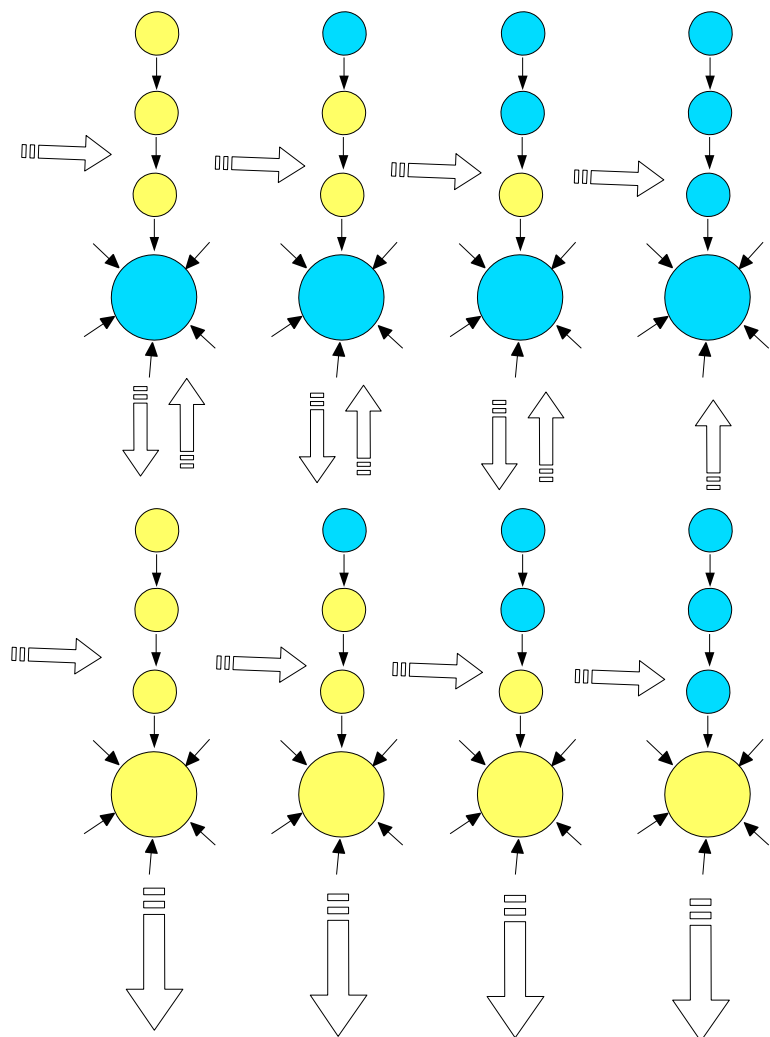
Here we calculate upper and lower bounds on the probability that a train of mutants, which arrives at the base of the stem with an initial length l , succeeds in taking over the root node and placing a new mutant in one of the reservoirs ($T = \mathbb{E}(l)$). We adapt the technique used in Díaz et al. [2013] and consider a finite state Markov process with two absorbing states: either a new mutant is placed in one reservoir, or the mutant train has disappeared. All other states correspond to a current train length and state of the root node. This is illustrated in Fig. 4.3.

For each state, the probability of eventually succeeding is $p_{i\uparrow}$, where $i \leq l$ indicates the current train length and \uparrow indicates whether the root node is occupied by a mutant, \uparrow , or resident, \downarrow . Clearly $p_{0\downarrow} = 0$ because the train has disappeared, the root is a resident and hence an absorbing state has been reached. Similarly, $p_{0\uparrow} = r/(B + r)$, denotes the probability that the mutant in the root node reproduces before being replaced by the offspring of residents in any of the B branches. By examining all possible transitions we obtain:

$$(B + r)p_{i\uparrow} = (B - 1)p_{i\downarrow} + r + p_{i-1\uparrow} \quad (4.8)$$

$$(r + 1)p_{i\downarrow} = rp_{i\uparrow} + p_{i-1\downarrow} \quad (4.9)$$

If the train is i mutants long, and the root is a mutant, the system can transition into the “successful” absorbing state, with relative weight r , can lose the root mutant, with relative weight $B - 1$ (because there are $B - 1$ other branches), or the train may erode, with relative weight 1, leaving the root node unchanged. If the root node is a resident, then the only possible actions are for the train to replace the root node, or for the train to be eroded from behind, with relatively probabilities r and 1 respectively.



Mutant Propagation

Figure 4.1: Possible states of the end of the stem, along with the transitions between them. Here mutants are yellow, residents blue. States further to the left, with longer mutant trains, have been omitted. We note the similarity to the argument of Díaz et al. [2013], but emphasize the simplicity gained by considering only the base of the stem, rather than the entire branch. This system contains far fewer states and far fewer transitions making it more amenable to analytic reasoning.

4.3. Interaction of trains with root node

Finally, coefficients on the left hand side of the equations normalise over all possible courses of action. Written as a matrix equation this gives:

$$\begin{bmatrix} B+r & 1-B \\ -r & r+1 \end{bmatrix} \begin{bmatrix} p_{i\uparrow} \\ p_{i\downarrow} \end{bmatrix} = \left(\begin{bmatrix} r \\ 0 \end{bmatrix} + \begin{bmatrix} p_{i-1\uparrow} \\ p_{i-1\downarrow} \end{bmatrix} \right)$$

which yields

$$\begin{bmatrix} p_{i\uparrow} \\ p_{i\downarrow} \end{bmatrix} = \frac{1}{B+2r+r^2} \begin{bmatrix} r+1 & B-1 \\ r & r+B \end{bmatrix} \left(\begin{bmatrix} r \\ 0 \end{bmatrix} + \begin{bmatrix} p_{i-1\uparrow} \\ p_{i-1\downarrow} \end{bmatrix} \right). \quad (4.10)$$

We now calculate both upper and lower bounds on the expected success probability, $\mathbb{E}(p_{i\downarrow})$, when our train first arrives at the root.

4.3.1 An upper bound on train success probability

By neglecting terms from the denominator of our fraction, and increasing several of our matrix entries, we find an upper bound on the R.H.S. of Eq. (4.10). This works because $p_{i\uparrow}, p_{i\downarrow} \geq 0$, and thus we are making the R.H.S. strictly more positive. It also significantly simplifies our equation.

$$\begin{bmatrix} p_{i\uparrow} \\ p_{i\downarrow} \end{bmatrix} < \frac{1}{B+2r+1} \begin{bmatrix} r+1 & B+r \\ r+1 & r+B \end{bmatrix} \left(\begin{bmatrix} r \\ 0 \end{bmatrix} + \begin{bmatrix} p_{i-1\uparrow} \\ p_{i-1\downarrow} \end{bmatrix} \right)$$

where the inequality applies element-wise. Substituting $p_{0\uparrow} = r/(B+r), p_{0\downarrow} = 0$ into the above gives

$$p_{1\uparrow} < \frac{r^2+r}{B+2r+1} \left(1 + \frac{1}{B+r} \right).$$

Using the fact that the upper bounds for $p_{i\uparrow}$ and $p_{i\downarrow}$ are equal (because both rows of the matrix are identical) gives:

$$p_{i\uparrow} < \frac{r^2+r}{B+2r+1} + \frac{(r+1)+(B+r)}{B+2r+1} p_{i-1\uparrow} = \frac{r^2+r}{B+2r+1} + p_{i-1\uparrow}.$$

By induction we find

$$p_{i\uparrow} < \frac{r^2+r}{B+2r+1} \left(i + \frac{1}{B+r} \right).$$

4.3. Interaction of trains with root node

This yields an upper bound for $p_{i\uparrow}$, which we then use in calculating a tighter bound for $p_{i\downarrow}$. From Eq. (4.8) and Eq. (4.9) we can derive:

$$\begin{aligned} p_{i\uparrow} &< \frac{Bp_{i\downarrow} + r + p_{i-1\uparrow}}{B} \\ (r+1)p_{i\downarrow} &< \frac{r}{B}(Bp_{i\downarrow} + r + p_{i-1\uparrow}) + p_{i-1\downarrow} \\ p_{i\downarrow} &< \frac{r}{B}(r + p_{i-1\uparrow}) + p_{i-1\downarrow} \end{aligned}$$

$$\Rightarrow p_{i\downarrow} < \sum_{n=1}^i \frac{r}{B}(r + p_{n-1\uparrow}) \leq \frac{ir^2}{B} \left(1 + \frac{r+1}{(B+2r+1)(B+r)} + \frac{H-1}{2} \frac{r+1}{B+2r+1} \right)$$

And thus we have $p_{i\downarrow} < ir^2(1 + \epsilon_{4+})/B$, where

$$\epsilon_{4+} = \frac{1+r}{B+2r+1} \left(\frac{1}{B+r} + \frac{H-1}{2} \right) \quad (4.11)$$

is our error term. This error term can be made arbitrarily small for sufficiently large B .

4.3.2 A lower bound on train success probability

For the lower bound, we must deal with the possibility that a small number δ of branches contain mutants. Because the train collision argument (section 4.2.2) is dependent on at most a single mutant existing in each branch, we wish to only consider reproductive events which place new mutants into branches that currently have no mutants, neglecting the rest (this causes no issue when calculating a *lower* bound). Further, trains from other mutants may compete for control of the root node, this must be accounted for. Although generically trains do not encounter one another, this lower bound is calculated as if all other mutant occupied branches have mutants at the base of their stems at all times. This arrangement, while unrealistic, describes the situation which minimizes the success probability of a given train, and is thus useful for finding lower bounds. The following equations are written under the assumption of this “worst case scenario” (worst from the perspective of the train we are focusing on).

$$\begin{bmatrix} B+r+(\delta-1)(r-1) & 1-B-(\delta-1)(r-1) \\ -r & r+1 \end{bmatrix} \begin{bmatrix} p_{i\uparrow} \\ p_{i\downarrow} \end{bmatrix} = \left(\begin{bmatrix} r \frac{B-\delta}{B} \\ 0 \end{bmatrix} + \begin{bmatrix} p_{i-1\uparrow} \\ p_{i-1\downarrow} \end{bmatrix} \right)$$

4.3. Interaction of trains with root node

$$\begin{bmatrix} p_{i\uparrow} \\ p_{i\downarrow} \end{bmatrix} = \frac{1}{B + \delta(r-1) + 1 + r + r^2} \begin{bmatrix} r+1 & B + \delta(r-1) - r \\ r & 1 + B + \delta(r-1) \end{bmatrix} \left(\begin{bmatrix} r \frac{B-\delta}{B} \\ 0 \end{bmatrix} + \begin{bmatrix} p_{i-1\uparrow} \\ p_{i-1\downarrow} \end{bmatrix} \right)$$

By ignoring the positive effects of $p_{i-1\uparrow}$ on $p_{i\downarrow}$ we form the inequality

$$p_{i\downarrow} > \frac{1}{B + \delta(r-1) + 1 + r + r^2} \left(r^2 \frac{B-\delta}{B} + (1 + B + \delta(r-1))p_{i-1\downarrow} \right)$$

for $i \geq 1$. By induction this leads to

$$\begin{aligned} p_{i\downarrow} &> \frac{B-\delta}{B} \frac{r^2}{B + \delta(r-1) + 1 + r + r^2} \sum_{n=0}^{i-1} \left(\frac{B+1+\delta(r-1)}{B + \delta(r-1) + r^2 + r + 1} \right)^n \\ &= \frac{B-\delta}{B} \frac{r^2}{B + \delta(r-1) + 1 + r + r^2} \frac{1 - \left(\frac{B+1+\delta(r-1)}{B + \delta(r-1) + r^2 + r + 1} \right)^i}{1 - \frac{B+1+\delta(r-1)}{B + \delta(r-1) + r^2 + r + 1}} \end{aligned}$$

By re-writing $\frac{B+1+\delta(r-1)}{B + \delta(r-1) + r^2 + r + 1}$ as $1 - \frac{r^2+r}{B + \delta(r-1) + r^2 + r + 1}$, and then simplifying the denominator of the equation we find

$$p_{i\downarrow} > \frac{B-\delta}{B} \frac{r^2}{r+r^2} \left(1 - \left(1 - \frac{r^2+r}{B + \delta(r-1) + r^2 + r + 1} \right)^i \right)$$

Using the binomial theorem to expand the inner bracket, we find an alternating series. As long as $i(r^2+r)/(B + \delta(r-1) + r^2 + r + 1) < 1$ the series expansion of the inner bracket will give an alternating sequence with monotone decreasing absolute terms. Truncating the series after three terms will preserve the inequality because the sum of the first three terms is greater than any subsequent sum. This leads to:

$$p_{i\downarrow} > \frac{B-\delta}{B} \frac{ir^2}{B + \delta(r-1) + r^2 + r + 1} \left(1 - \frac{i-1}{2} \frac{r^2+r}{B + \delta(r-1) + r^2 + r + 1} \right).$$

Remembering that $1/(1+x) < 1-x$ whenever $x > -1$, we find:

$$p_{i\downarrow} > \frac{ir^2}{B} \left(1 - \frac{\delta}{B} - \frac{\delta(r-1) + r^2 + r + 1}{B} - \frac{H-1}{2} \frac{r^2+r}{B} + O(B^{-2}) \right)$$

We are free to drop the very small positive terms at the end, and find the result $p_{i\downarrow} > (1 - \epsilon_{4-})ir^2/B$, where

$$\epsilon_{4-} = \frac{\delta r + r^2 + r + 1}{B} + \frac{(H-1)(r^2+r)}{2B} \quad (4.12)$$

can be made small whenever $\delta, H \ll B$. Thus $\mathbb{E}(p_{l\downarrow}) > Tr^2B^{-1}(1 - \epsilon_{4-})$. Armed with the constraints

$$(1 - \epsilon_{4-})ir^2/B < p_{i\downarrow} < (1 + \epsilon_{4+})ir^2/B, \quad (4.13)$$

we note that $\mathbb{E}(p_{l\downarrow}) \approx \mathbb{E}(ir^2/B) = Tr^2/B$.

4.4 Bounds on fixation probabilities

4.4.1 Upper bound

The probability of transitioning from 1 to 2 mutants in the reservoir serves as an upper bound on the mutant fixation probability. We therefore make several optimistic (from the mutants point of view) assumptions:

1. The original mutant appears in a reservoir node (ignoring ϵ_0 , see Eq. (4.1)).
2. Simplified train launching probability (ignoring ϵ_1 , see Eq. (4.2)).
3. No detrimental effects based on our initial conditions (ignoring ϵ_2 , see Eq. (4.3)).
4. No train collisions (ignoring ϵ_3 , see Eq. (4.7))
5. We use the upper bound for the probability that a train succeeds in producing another reservoir mutant ($Tr^2(1 + \epsilon_{4+})/B$, see Eq. (4.11)).

A single mutant in any reservoir produces a new train with a probability of at most $r^2/(F_tL)$ per time step. Subsequently, each train succeeds in placing another mutant in any reservoir with a probability of at most $(1 + \epsilon_{4+})Tr^2/B$. At the same time, the central root node has a probability of at least $\frac{B-1}{B+r-1} \frac{1}{F_tBL}$ to remove the mutant node from the reservoir. Thus, the chance of the mutant producing a successful train before being erased by the root node is *at most*:

$$\rho_H \leq \rho_{H+} = \frac{Tr^4(1 + \epsilon_{4+})}{Tr^4(1 + \epsilon_{4+}) + \frac{B-1}{B+r-1}} \approx 1 - \frac{1}{Tr^4 + 1} \quad (4.14)$$

The approximation in Eq. (4.14) becomes exact in the limit of large B . Note that because of $T < H$ we can replace T by H in Eq. (4.14) to obtain a more generous upper bound.

4.4.2 Lower bound

On the slow timescale the dynamics can be approximated by a random walk X_t on the number of mutants in the reservoir. When reservoir mutants are rare, our forward bias is as calculated previously. Unfortunately, because the analytic arguments in the previous sections are based on the assumption of having only a few reservoir mutants, further arguments must be made to determine a lower bound on forward bias when reservoir mutants are common.

All changes made in the reservoir can ultimately be treated as the descendants of some reservoir occupant replacing the occupant of some other reservoir node. The chance of *any particular* reservoir mutant replacing *any particular* reservoir resident is *always* higher than the converse, as even in the worst case scenario where a mutant is competing with $L - 1$ other mutants in its branch, and a reservoir resident suffers from no such competition (thus eradicating and advantage the mutant might attain along the stem), the mutant will still have an advantage of at least r when competing for and propagating from the central root node. This would indicate that, conditioning on a particular pair of nodes replacing one another, the bias must be at least r^2 . Because all pairwise interactions between reservoir nodes are biased in the mutants favour, the random walk itself must be at least somewhat biased in the mutants favour for all values of X_t . A random walk on the integers from 0 to $BL + 1$ with forward bias γ for $X_t < \delta \ll B$ and no bias for $X_t \geq \delta$ will thus significantly underestimate the fixation probability of the true process. We bound our walk at $BL + 1$ rather than BL because we require not only that all reservoir nodes are mutants, but also that this propagates down to the root and all stem nodes. By overshooting our intended target we demand that the system remain in the state BL long enough to propagate forward- and hence long enough that all stems will have been replaced by mutant reservoir nodes multiple times (as changes to the reservoir nodes are on a slower timescale). Please note that the actual superstar never enters a state with $BL + 1$ reservoir mutants, and this is merely a useful tool for dealing with the random walk. The fixation probability of the random walk described thus acts as a lower bound on the fixation probability of the true process. In the following, we assume that H and r are fixed and that $B, L \gg H$.

In order to determine the fixation probability of the random walk X_t , we construct a martingale, $Q(X_t)$. A martingale is a stochastic process such

4.4. Bounds on fixation probabilities

that the expected value in the next time step is equal to the current value:

$$\mathbb{E}(Q(X_{t+1})|X_1, X_2 \dots X_t) = Q(X_t). \quad (4.15)$$

Because our system is Markov, it is sufficient to condition only on X_t . For $Q(X_t)$ to be a martingale, we thus require:

$$Q(k) = \begin{cases} \frac{\gamma}{1+\gamma}Q(k+1) + \frac{1}{1+\gamma}Q(k-1) & 0 < k < \delta & \text{(forward bias)} \\ \frac{1}{2}Q(k+1) + \frac{1}{2}Q(k-1) & \delta \leq k < BL+1 & \text{(no bias)}. \end{cases} \quad (4.16)$$

These constraints admit the solution $Q(k) = \gamma^{-k}$ for $k < \delta$, and $Q(k) = Ak + D$ for $k \geq \delta$. For $Q(k)$ to satisfy the martingale conditions as needed, we demand $\delta \in \mathbb{N}$. The constants A, D are determined by connecting the solutions for the two regions. In particular,

$$\gamma^{-\delta} = Q(\delta) = A\delta + D$$

must hold such that $Q(\delta)$ is well defined and

$$2(A\delta + D) = \gamma^{-\delta+1} + A\delta + A + D$$

to satisfy the martingale property at δ . Thus,

$$\begin{aligned} A &= \gamma^{-\delta}(1 - \gamma) \\ D &= \gamma^{-\delta}(1 - \delta(1 - \gamma)) \end{aligned}$$

which yields

$$\begin{aligned} Q(0) &= 1 \\ Q(1) &= \gamma^{-1} \\ Q(BL+1) &= \gamma^{-\delta} + \gamma^{-\delta}(1 - \gamma)(BL + 1 - \delta). \end{aligned}$$

Let τ be the first time X_t reaches one end of the random walk. Because $Q(k)$ is bounded for all relevant values of k we are able to invoke the optional stopping theorem [Klenke, 2006, p. 210]. The O.S.T. then implies that

$$Q(1) = Q(X_0) = \mathbb{E}(Q(X_\tau)) = Q(0)P(0) + Q(BL)P(BL),$$

4.4. Bounds on fixation probabilities

where $P(0)$ and $P(BL)$ represent the probabilities of reaching either end of our random walk. Using $P(0) = 1 - P(BL)$ we find

$$P(BL) = \frac{Q(1) - Q(0)}{Q(BL) - Q(0)} = \frac{1 - \gamma^{-1}}{1 - \gamma^{-\delta} - \gamma^{-\delta}(1 - \gamma)(BL + 1 - \delta)}. \quad (4.17)$$

In order to keep error terms small, we must select δ such that $\gamma^\delta \gg BL$ and $\delta \ll B, L$. As long as B and L are large and sufficiently similar, this is possible provided that $\gamma > 1$ (to be shown). Thus, for any choice of H , and for any $r > 1$ we can select B and L such that

$$P(BL) = \frac{Q(1) - Q(0)}{Q(BL) - Q(0)} = \frac{1 - \gamma^{-1}}{1 + \epsilon_5} \quad (4.18)$$

with

$$\epsilon_5 = \gamma^{-\delta} ((\gamma - 1)(BL + 1 - \delta) - 1) \ll 1.$$

In order to bound ϵ_5 , a lower bound on γ is required. This is found by taking the lower bound on the production rate of successful trains, $(1 - \epsilon_1)(1 - \epsilon_3)(1 - \epsilon_{4-})Tr^4/(BLF_T)$, and comparing to our upper bound on the removal probability for reservoir mutants, $1/(BLF_t)$. This represents the eventual forward bias once the first node in the stem has been replaced at least once. To account for the possibility of mutant loss before the first stem node is replaced we must consider ϵ_2 , which acts as an additive penalty (because it only applies once per reservoir mutant). This yields

$$\gamma \geq r^4 T (1 - \epsilon_1)(1 - \epsilon_3)(1 - \epsilon_{4-}) - \epsilon_2. \quad (4.19)$$

In the limit of large B and L all error terms tend to zero. Thus, to show $\gamma > 1$, it is sufficient to show that $r^4 T > 1$. Recalling that $A_t - Z_t$ represents the length of a train at time t (see 4.2), and noting that it is submartingale (the expected future value is greater than the current value) whenever $r > 1$, we can easily show that $T = \mathbb{E}(A_\tau - Z_\tau) \geq A_0 - Z_0 = 1$, and thus, in the limit, $\gamma \geq r^4 > 1$. Thus ϵ_5 can be made arbitrarily small.

Substituting Eq. (4.19), our lower bound on γ into Eq. (4.18) yields a lower bound on $P(BL)$, which in turn provides a lower bound on the fixation probability.

In cases where generic bounds not dependent on T are desired, the lower bound $(H - 1)(1 - 1/r)^2$ can be substituted into Eq. (4.19) in T 's place, resulting in a looser bound on fixation probability.

4.5 Bringing it all together

Collating Eq. (4.18) and Eq. (4.14), we find:

$$\frac{1 - \epsilon_0}{1 + \epsilon_5} \left(1 - \frac{1}{r^4 T (1 - \epsilon_1)(1 - \epsilon_3)(1 - \epsilon_{4-}) - \epsilon_2} \right) \leq \rho_H \leq 1 - \frac{\frac{B-1}{B+r-1}}{Tr^4(1 + \epsilon_{4+}) + \frac{B-1}{B+r-1}} \quad (4.20)$$

where

$$T = (1 - \alpha)^{H-2} \sum_{z=1}^{H-1} (H - z) \alpha^{z-1} \left[\binom{H - 3 + z - 1}{z - 1} - \binom{H - 3 + z - 1}{z - 2} \right],$$

$$(H - 1)(1 - r^{-1})^2 \leq T \leq H, \text{ Length of train in stem, Eq. (3.2), section 4.2.1}$$

$$\epsilon_0 = \frac{1 + HB}{BL + 1 + HB}, \text{ Chance that initial mutant is not in reservoir, Eq. (4.1)}$$

$$\epsilon_1 = \frac{r - 1}{L + r - 1}, \text{ Simplification of train launch probability, Eq. (4.2)}$$

$$\epsilon_2 = \frac{1}{1 + BL^2}, \text{ Mutant removal before top of stem updates, Eq. (4.3)}$$

$$\epsilon_3 = H \frac{r^2}{L + r - 1 + r^2}, \text{ Upper bound on train collisions, Eq. (4.7)}$$

$$\epsilon_{4-} = \frac{\delta r + r^2 + r + 1}{B} + \frac{(H - 1)(r^2 + r)}{2B}, \text{ Train success, lower bound Eq. (4.12)}$$

$$\epsilon_{4+} = \frac{(r + 1)}{B + 2r + 1} \left(\frac{1}{B + 1} + \frac{H - 1}{2} \right), \text{ Train success, upper bound Eq. (4.11)}$$

$$\epsilon_5 = \gamma^{-\delta} [(\gamma - 1)(BL + 1 - \delta) - 1], \text{ Martingale error term, Eq. (4.18) .}$$

All error terms can be simultaneously made small when $B, L \gg H, \delta$ and $(r^4 T)^\delta \gg BL$. To find looser upper and lower bounds independent of T we can substitute in our upper and lower bounds for T (respectively) into Eq. (4.20).

In the particular limit $B \rightarrow \infty$, with $\sqrt{B} - 1 < \delta \leq \sqrt{B}$ and $L = B$ all error terms disappear. Other L, B relations are possible – for example $L = B^\beta$ will also work if $\beta > 0$, although $L = \beta^B$ is liable to cause problems for our ϵ_5 bound, as it is not immediately clear that $\gamma^{-\delta} \beta^B$ must tend to zero in the limit.

Even though fixation probabilities can be made arbitrarily close to 1 on large superstars and sufficiently large H , the fixation probability remains bounded away from 1 for any finite graph. As a concrete example, consider

4.5. Bringing it all together

$r = 2$ and $H = 50$, which leads to $T \approx 13.25$ and $0.995283 \leq \rho_{50} \leq 0.995306$ in the appropriate limits of large B, L . In comparison, a sizable superstar with $B = L = 5000$ ($N \approx 2.5 \cdot 10^5$) yields $0.985323 \leq \rho_{50}^N \leq 0.995375$, which includes all error terms. The fixation probability for a similarly sized isothermal graph is just short of 0.5. It is interesting to note that in this case, the greatest source of error (in the lower bound) is the possibility that the original mutant will arise in either the stem or root, leading to fixation with low probability.

Chapter 5

In which some loose ends are tied up and others are left open

It turns out that superstars act as evolutionary amplifiers only under very specific conditions. Here we discuss the most important requirements.

5.1 Selection & sequence

The original Moran process is formulated as a fecundity based birth-death process, that is, fitness affects the rate of birth (reproduction) whereas death (replacement) occurs uniformly at random. Alternatively, fitness could just as well affect survival such that birth events occur uniformly at random but death events occurring with probability inversely proportional to fitness. Similarly, the sequence of events could be reversed such that first an individual dies and then the remaining individuals compete to repopulate the vacant site. This yields a total of four distinct scenarios: Bd , bD , dB and Db , where capital letters refer to the fitness dependent selection step. The original Moran process corresponds to Bd and the fixation probability is given in Eq. (1.1). In unstructured populations the four dynamical scenarios result in only marginal differences in fixation probabilities. However, they can have crucial effects on the evolutionary outcome in structured populations [Ohtsuki et al., 2006, Ohtsuki and Nowak, 2006, Zukewich et al., 2013]. Frean and Baxter [2008] examine all four cases for both complete graphs, and star graphs, showing that stars act as evolutionary suppressors in both the dB and Db cases, and are significantly less effective in the bD case compared to the original Bd case. Similar results apply to superstars:

bD updates: For the birth-death process with selection on survival, mutants only gain any advantage when the root node reproduces. Whenever any other node reproduces, there is only a single downstream node, and thus

no opportunity for competition, rendering any fitness advantage irrelevant. This lack of advantage in the stem leads to an expected train length of 1, regardless of stem length or mutant's fitness. The chance of launching a successful train in a given time step is $1/(NLB)$ and the chance of replacing the original mutant is $1/(NBLr - N(r - 1))$. This results in a bias of approximately $\frac{r}{1+r}$ for the initial mutant to eventually create a second reservoir mutant – the same bias as for the original Moran process. Thus, we might expect fixation probabilities similar to the original Moran process on BL nodes, and certainly nowhere close to the amplification observed for Bd .

Db updates: For the death-birth process with selection on survival, the prospects of mutants drop even further. The probability to successfully place even a single offspring in the top of the stem is only $r/(L + r)$. As the train propagates along the stem it tends to grow because the mutant at the back of the train is less likely to die than the resident in front of it, leading to the same train dynamics observed for the Bd process. Note that for death-birth processes the top of the stem no longer changes on the fast timescale and hence trains start at the top instead of the second node. Upon reaching the end of the stem the train competes with the other branches for control over the root node and succeeds with probability near rT/B (over the lifetime of the train). Once a mutant occupies the root, it is predestined to have many offspring – in each time step a reservoir node dies with high probability and gets replaced by an offspring of the mutant in the root node, whereas the probability is low that the root node is replaced. More specifically, we expect $rN/(1 + r)$ reservoir nodes to become mutants before the root node is replaced. At that point it is reasonable to assume that mutants reach fixation with high probability. We conjecture that the probability of mutant fixation on the superstar is close to the probability of a mutant eventually being placed in the root node. Thus, we expect a fixation probability close to r^2T/BL . This result is significantly *less* than the almost certain fixation found in superstars for the Bd process, and is in fact smaller than even the fixation probability of the original Moran process (c.f. Eq. (1.1)), suggesting that for Db dynamics superstars act as strong evolutionary suppressors. The result does, however, match well with the $1/N$ scaling found for the fixation on stars [Freen and Baxter, 2008].

dB updates: The final case is the death-birth process with selection on reproduction. Once again the probability of placing a mutant offspring in the

stem before losing the reservoir mutant is near r/L , but now without further benefits along the stem. Consequently, trains that do reach the root still have an expected length of 1. Thus, each train has a probability of roughly r/B for claiming the root node, which then produces $N/2$ mutants in the reservoirs, on average – enough to suggest fixation with high probability, but less than for Db . Thus, we conjecture fixation probabilities near r^2/N – the worst outcome of the four scenarios. Once again, we note the significant penalty as compared to the original Moran process as well as the similarities to the $1/N$ scaling for stars [Freen and Baxter, 2008].

5.2 Mutations

Even though we did not explicitly model the process of mutation, we implicitly assumed that mutations are rare and arise spontaneously in any node selected uniformly at random. For the superstar this means that most mutations arise in a reservoir node – simply because the overwhelming majority of nodes are reservoir nodes.

An alternative and equally natural assumption is that mutations arise during reproduction events. Such a change does not affect the fixation probabilities in the original Moran process. However, in highly heterogeneous population structures crucial differences in the fixation probabilities can arise because mutants preferentially arise in certain locations [Maciejewski et al., 2014]. For superstars, when using the Bd or bD update rules, mutants most likely arise at the top of a stem. This is an unfortunate position because the mutant is likely to be replaced before reproducing even once – extinction is almost certain. In contrast, for Db and dB , mutants again most likely arise among the reservoir nodes – but for those updates superstars do not act as evolutionary amplifiers.

Even though the dynamical properties of superstars are intriguing, the list of caveats demonstrates that the evolutionary amplification is highly sensitive to the details of the model – maybe this is the reason that superstar-like structures have not been reported in nature.

5.3 Deleterious mutations, $r < 1$

As a final remark, we note that all discussion so far is based on the assumption of a beneficial mutation, $r > 1$. In addition to promoting beneficial mutations, an evolutionary amplifier must also suppress the fixation of deleterious mutations, $r < 1$. Here we argue that a deleterious mutant indeed

disappears almost surely on sufficiently large superstars.

Consider a single mutant with fitness 1, in a population of residents with fitness $1/r$. Note that we can rescale “fitness” without changing the dynamics of the system, since fitness is used only as a weighting factor and never in an absolute sense, only relative fitness matters. We next observe that all calculations performed previously with respect to a rare mutant with a fitness advantage would now apply to a resident, if it were to become rare - that is, if residents were rare we would expect trains of residents to propagate down the stem, incrementing with probability $1/(r+1) > 1/2$ and shrinking with probability $\alpha = r/(1+r) < 1/2$, leading to $T > (H-1)(1-r)^2$. The same martingale argument that we previously used to find a lower bound on mutant fixation probability can now be used to form a lower bound on resident fixation probability. This time we use X_t to track the number of residents in our reservoir nodes. Thus $X_0 = BL - 1$. X_t can be treated as a random walk that has forward bias $\gamma \approx r^{-4}T$ when residents are rare, and no bias otherwise. This leads to a formula for the fixation probability of *residents*

$$\begin{aligned}
 1 - \rho_H &\geq P(BL) = \frac{Q(0) - Q(BL - 1)}{Q(0) - Q(BL)} \\
 &= \frac{1 - \gamma^{-\delta} - \gamma^{-\delta}(1 - \gamma)(BL - 1 - \delta)}{1 - \gamma^{-\delta} - \gamma^{-\delta}(1 - \gamma)(BL - \delta)} \\
 &= 1 - \frac{\gamma^{-\delta}(\gamma - 1)}{1 - \gamma^{-\delta} - \gamma^{-\delta}(1 - \gamma)(BL - \delta)} \\
 \rho_H &\leq \frac{\gamma^{-\delta}(\gamma - 1)}{1 - \gamma^{-\delta} - \gamma^{-\delta}(1 - \gamma)(BL - \delta)} \\
 \rho_H &\leq \gamma^{1-\delta}
 \end{aligned}$$

In the above we require γ to be the bias in favour of the *resident*. In order to calculate it we need to find the expected train lengths of our resident trains. Structurally the train equation is the same as previously, but with r replaced by $1/r$ (the fitness of our resident strain). Because $1/r$ is now greater than one, all train length arguments that previously depended on $r > 1$ for beneficial mutations can now be applied to the resident when $1/r > 1$. Thus, we can show that the expected length of resident trains is large. Because $\gamma \approx r^{-4}T$, large T leads directly to a large γ (in the residents favour), and $\gamma^{-\delta}BL \ll 1$. Our Martingale will significantly underestimate the fixation probability for our resident, thus we expect the above bound to overestimate the probability of mutant fixation. In order to calculate exact bounds on our new γ we must apply error terms similar to $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_{4-}$,

5.3. Deleterious mutations, $r < 1$

giving us lower bounds on the effectiveness of “resident trains”. No error term equivalent to ϵ_0 will be required, as the possibility of our initial mutant being placed not in a reservoir node can only reduce our mutant fixation probability. Thus, it can be seen that the fixation probability of a deleterious mutant tends to zero, and in particular that $\rho_H \leq \gamma^{1-\delta}$. In the limit as B and L are taken to infinity (in an appropriate manner) we find $\rho_H \leq (r^{-4}T)^{1-\delta}$

Chapter 6

In which concluding remarks occur

In this thesis I have explained and explored two mutually contradictory proofs. In doing so, I went beyond the counter-example provided by Díaz et al. [2013], and managed to identify the errors in Lieberman et al. [2005] that led the argument astray, thus plugging a distinctive gap in the literature, and providing an *explanation* for the contradiction between the two proofs, rather than merely a proof that one of them was wrong.

Further than that, I was able to determine true limits on the fixation probability of superstars.

In completing this proof I have drawn upon a wide variety of mathematical and computational techniques, ranging from basic simulations, induction, and a wide range of algebraic identities, to more conceptual tools such as Martingales and reflection principles.

While this paper may have answered at least one question regarding superstars- namely the identification of the true fixation probability – many questions still remain. While chapter 5 gives some suggestions regarding how to proceed in finding the fixation probability under a variety of different regimes, it by no means provides solid proof. Lieberman et al. [2005] states that the fixation probability for both funnels and metafunnels are also governed by Eq. (1.3), but given the difficulties regarding superstars, it would seem prudent for such claims to be investigated further. One might also question the interaction between the superstar structures and fixation *time* – in pushing the fixation probability towards certainty, do we also push the fixation time to infinity, placing our so called certainty out of reach?

Superstars represent the most prominent representatives of evolutionary amplifiers – structures that are capable of increasing selection and suppressing random drift. For $r > 1$ and in the limit of large N we have derived upper and lower bounds for the fixation probability, ρ_H :

$$1 - \frac{1}{r^4 T} \leq \rho_H \leq 1 - \frac{1}{1 + r^4 T}, \quad (6.1)$$

where $(H - 1)(1 - 1/r)^2 \leq T \leq H$.

The upper bound for ρ_H in Eq. (6.1) results in a contradiction with the originally reported fixation probability, Eq. (1.3), for sufficiently large r . For the specific case of $H = 3$ the discrepancy was pointed out in Díaz et al. [2013]. At the same time, the lower bound for ρ_H confirms that superstars are indeed capable of providing an arbitrarily strong evolutionary advantage to any beneficial mutation, as suggested in Lieberman et al. [2005]. Using symmetry arguments, it also follows that for $r < 1$ the fixation probability can be made arbitrarily small (see 5.3).

In the case $H = 2$ (or $k = 4$ in Lieberman et al. [2005]) we obtain an expected train length of $T = 1$ and recover the original bias, $\frac{r^4}{1+r^4}$. Discrepancies arise only for $H \geq 3$ (or $k \geq 5$) but those cases were not included in the simulations in Lieberman et al. [2005]. For $H = 3$, we obtain $T = 2r/(1+r)$, which results in a bias of $\frac{2r^5}{1+r+2r^5}$ and recovers the upper bound reported by Díaz et al. [2013], but our derivation offers a more illuminating explanation. The agreement extends to higher values of H when extending the technique in Díaz et al. [2013] numerically (using the program available in A.1).

An appropriately skeptical reader might ask why the theory presented here should be trusted over those previously presented in the literature – after all, both claim to offer rigorous proof. First, we note the agreement between predictions made here, and both Lieberman et al. [2005] and Díaz et al. [2013] for the appropriate values of H . Second, we identify correlations between neighbouring stem nodes as the cause for the discrepancies between the two previous papers. Finally, we invite readers to scrutinize the proof offered here most thoroughly, to convince themselves of its rigor. Superstars have already presented unexpected subtleties, and as always, we need caution and vigilance to discern between scientific selection and random drift.

Bibliography

- T. Antal, S. Redner, and V. Sood. Evolutionary dynamics on degree-heterogeneous graphs. *Physical Review Letters*, 96(18):188104, 2006.
- M. Broom and J. Rychtář. An analysis of the fixation probability of a mutant on special classes of non-directed graphs. *Proceedings of the Royal Society A*, 464:2609–2627, 2008. doi: 10.1098/rspa.2008.0058.
- R Bürger and R Lande. On the distribution of the mean and variance of a quantitative trait under mutation-selection-drift balance. *Genetics*, 138, 1994.
- Alex Murray Burton Voorhees. Fixation probabilities for simple digraphs. *Proc. R. Soc. A*, 469, 2013.
- Josep Díaz, Leslie Ann Goldberg AD George B. Mertzios, David Richerby, Maria Serna, and Paul G. Spirakis. On the fixation probability of superstars. *Proc. R. Soc. A*, 469, 2013.
- M. Frean, P. Rainey, and A. Traulsen. The effect of population structure on the rate of evolution. *Proceedings of the Royal Society B*, 280(1762): 20130211, 2013. doi: 10.1098/rspb.2013.0211.
- M. R. Frean and G. J. Baxter. Death-birth ordering and suppression of fitness in networks. working paper: <http://homepages.mcs.vuw.ac.nz/~marcus/manuscripts/FreanBaxterJTB.pdf>, October 2008.
- F. Fu, M. A. Nowak, and C. Hauert. Evolutionary dynamics on graphs: Efficient method for weak selection. *Physical Review E*, 79:046707, 2009.
- Feng Fu and Martin A Nowak. Global migration can lead to stronger spatial selection than local migration. *Journal of Statistical Physics*, 151:637–653, 2013. doi: 10.1007/s10955-012-0631-6.

- Margaret J. Eppstein Joshua L. Payne. Evolutionary dynamics on scale-free interaction networks. *IEEE, transactions on evolutionary computation*, 2009.
- Achim Klenke. *Probability Theory: A comprehensice Course*. Springer, 2006.
- Leonid B. Koralov and Yakov G. Sinai. *Theory of Probability and Random Process*. Springer, 2nd edition, 2007.
- E. Lieberman. *Evolution and the emergence of structure*. Harvard University Press, 2010.
- E. Lieberman, C. Hauert, and M. A. Nowak. Evolutionary dynamics on graphs. *Nature*, 433:312–316, 2005. ISSN 1476-4687.
- W. Maciejewski, F. Fu, and C. Hauert. Evolutionary game dynamics in populations with heterogeneous structures. *PLoS Computational Biology*, 10(4):e1003567, 2014.
- T. Maruyama. Effective number of alleles in a subdivided population. *Theoretical Population Biology*, 1:273–306, 1970.
- F. Michor, Y. Iwasa, and M. A. Nowak. Dynamics of cancer progression. *Nature Reviews Cancer*, 4:197–205, 2004.
- P. A. P. Moran. *The Statistical Processes of Evolutionary Theory*. Clarendon Press, Oxford, 1962.
- M. A. Nowak. *Evolutionary Dynamics*. Harvard University Press, Cambridge MA, 2006.
- M. A. Nowak and R. M. May. Evolutionary games and spatial chaos. *Nature*, 359:826–829, 1992.
- M. A. Nowak, F. Michor, and Y. Iwasa. The linear process of somatic evolution. *Proceedings of the National Academy of Sciences USA*, 100:14966–14969, 2003.
- M. A. Nowak, A. Sasaki, C. Taylor, and D. Fudenberg. Emergence of cooperation and evolutionary stability in finite populations. *Nature*, 428:646–650, 2004.
- H. Ohtsuki and M. A. Nowak. The replicator equation on graphs. *Journal of Theoretical Biology*, 243:86–97, 2006.

- H. Ohtsuki, C. Hauert, E. Lieberman, and M. A. Nowak. A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, 441: 502–505, 2006.
- Paulo Shakarian and Patrick Roos. Fast and deterministic computation of fixation probability in evolutionary graphs. In T. Znati, Z. Bar-Joseph, and L. Rothrock, editors, *Computational Intelligence and Bioinformatics / Modelling, Identification, and Simulation*, Calgary AB, 2011. ACTA Press. URL <http://shakarian.net/papers/SHAKRIAN-ROOS-fixation-prob-alg-CIB11.pdf>.
- M. Slatkin. Fixation probabilities and fixation times in a subdivided population. *Evolution*, 35:477–488, 1981.
- Joshua Zukewich, Venu Kurella, Michael Doebeli, and Christoph Hauert. Consolidating birth-death and death-birth processes in structured populations. *PLoS One*, 8(1):e54639, 2013.

Appendix A

Programs

A.1 superstarSolve

SuperstarSolve.m is a piece of matlab code I used to test various predictions for the probability of early extinction against the numerical results found using markov chains on possible states of a branch. The code requires a particular k, r, L and B , and hence can not be used in a general sense. $k = H + 2$. This is a polished version of the code, with slight changes in output format and variable names compared to the original, but is effectively identical.

This program can be found online at:

<https://github.com/alastair-JL/Academic/tree/master/Superstar-Project-2013-2014> .

```
function [x,b,A] = superstarSolve(k,r,L,B)
A=zeros(2^k,2^k);
b=zeros(2^k,1);
q=zeros(k,1);
Row=1;
A(1,1)=1;

while (Row<2^k)
q(1)=q(1)+1;
Row=Row+1;
jjj=1;
while (jjj<k)
if (q(jjj)>1.2)
q(jjj)=0;
q(jjj+1)=1+q(jjj+1);
end
jjj=jjj+1;
end
%%have now calibrated the q to the correct label.
```

```

if (q(k)>0)
b(Row)=-r;
if(q(k-1)==1)
A(Row,Row-2^(k-1))=L-1;
end
if(q(k-1)==0)
A(Row,Row-2^(k-1))=L;
end
end
if (q(k)==0 && q(1)==1)
A(Row,Row-1)=1/(L*B);
end % Possible actions of the center,
%either ending the test, or erasing X
jjj=1;
while (jjj<k-1)
jjj=jjj+1;
if (q(jjj)==1 && q(jjj+1)==0)
A(Row,Row+2^(jjj))=r;
end
if (q(jjj)==0 && q(jjj+1)==1 && jjj+1~k )
A(Row,Row-2^(jjj))=1;
end
end %%mutation passing down chain
%%special behavior of first node.
if (q(1)==1 && q(2)==0)
A(Row,Row+2)=r;
end
if (q(1)==1 && q(2)==1)
A(Row,Row-2)=B-1;
end
if (q(1)==0 && q(2)==1)
A(Row,Row-2)=B;
end
jjj=0;%%sum up rows to get diagonals...
sum=0;
while (jjj<2^k)
jjj=jjj+1;
sum= sum + A(Row,jjj);
end
sum=sum- b(Row);

```

A.1. superstarSolve

```
A(Row,Row)= - sum;
end

x=A\b;

PropagationProbability= x(2)

LiebermanPrediction= r^k/(1+r^k)

LieberRatio= (1-PropagationProbability)/(1-LiebermanPrediction)

if(k==5)
DiazPrediction= 2*r^5/(1+r+2*r^5)
DiazRatio= (1-PropagationProbability)/(1-DiazPrediction)
end

if(k>=4&& k<=10)
    %%Note, Train length makes no predictions for small k, and quickly
    %%breaks the machine for large k.
T= makeTrainLength(k,r);
AJLpredict= r^4*T/(r^4*T+1)
AJLRatio= (1-PropagationProbability)/(1-AJLpredict)
end

end

%%A few comments to those who may use this program in the future.

%% I used the ratio of the FAILURE PROBABILITIES compared
%% to the actual result for the failure probabilities as my
%% test. This is because failure probability is close to zero,
%% and thus its ratio is more sensitive then dividing a
%% bunch of things near one.

%%It is worth noting that for k=5, Diaz's predictions get more and more
%%accurate for larger values of L(leaf number) and B(branch number).
%%Lieberman's predictions plateau- reaching a maximal accuracy for
%% any given k,r combination.

function [T] = makeTrainLength(k,r)
```

```
    if (k==4)
        T=1;
    else
        H=k-2;
    alpha= 1/(1+r)
    T=(H-1)*alpha^0*nchoosek(H-4+1,1-1);
    z=2;
    while(z<H)
        T=T+(H-z)*alpha^(z-1)*(nchoosek(H-4+z,z-1)-nchoosek(H-4+z,z-2) );
        z=z+1;
    end
    T=(r/(1+r))^(H-2)*T;
end

end
```

A.2 SimpleChainTest

SimpleChainTest is a Java program that simulates a single chain for a large amount of time, and was used for the simulations in this paper. This version is an extension of the version used for my initial investigation, and collects significantly more data than the original. The underlying simulation structure remains unchanged.

This program can be found online at:

<https://github.com/alastair-JL/Academic/tree/master/Superstar-Project-2013-2014> .

```
package simplestartest;

import java.util.Random;

/**
 *
 * @author Babbelfish
 */
public class SimpleChainTest {

    /**
     * @param args the command line arguments
     */
}
```

A.2. SimpleChainTest

```
public static void main(String[] args) {

    int k=15;
    boolean weirdRecord=true;
    boolean changeMade=false;
    int numsuccess=0;
    double Mcount[]= new double[k];
    int MM[]= new int[k-1];
    int TrainArrive[][]= new int[k][k];
        //first index refers to location, second to train length.
    int TrainTime[][]= new int[k][k];
        //first index refers to location, second to train length.
    int MostRecentTrainLength[]=new int[k];
        // Index refers to location;
    long WW[]= new long[k-1];
    int numTrainsLeave=0;
    long numsample=0;
    int printseperator=4;
    double TotalTotalWeight=0;
    String trainPrint;

    for (int jij=0; jij<k-1; jij++){
        Mcount[jij]=0;
        MM[jij]=0;
        WW[jij]=0;
    }
    Mcount[k-1]=0;

    for (int iii=0; iii<20; ++iii){

        boolean chain[]= new boolean[k];
        double r=1.5;
        Random picker= new Random();

        chain[0]=true;
        int M=20000;
        double totalweight= chain.length;
        double pick=0;
```


A.2. SimpleChainTest

```
boolean success=false;
boolean extinct=false;
while (numsample< Long.MAX_VALUE-5){
    changeMade=false;

    chain[0]= (picker.nextDouble()*(M +r-1)<r);
    totalweight=0;
    for (boolean b: chain){
        totalweight+= (b?r:1);
    }
    TotalTotalWeight+=totalweight;

    pick= picker.nextDouble()*totalweight;
    int select=0;
    double pickcopy = pick;

    while (pickcopy>0 && select<k){
        pickcopy-= (chain[select]?r:1);
        select++;
    }

    select--;

    if (select==k-1){
        changeMade=true;
        //if final rung... do nothing!
    }else{
        if (chain[select] && !chain[select+1]){ //new train arrive
            int trainLength=0;
            while((trainLength<select)&& (chain[select-trainLength]) ) {
                trainLength++;
            }
            TrainArrive[select+1][trainLength]++;
            MostRecentTrainLength[select+1]=trainLength;
        }
        if (chain[select+1]!=chain[select]){
            changeMade=true;
        }
        chain[select+1]=chain[select];

        if (chain[select]&& select==0){
```

A.2. SimpleChainTest

```
        numTrainsLeave++;    }

    }

    for (int ppp=0; ppp<k;ppp++){
//loop to add time to correct train length.
        if (chain[ppp]){
            TrainTime[ppp][MostRecentTrainLength[ppp]]++;
        }
    }

    if (changeMade || !weirdRecord){
for (int jij=0; jij<k; ++jij){
    Mcount[jij]+= (chain[jij]?1:0);
    if (jij<k-1 && chain[jij] &&chain[jij+1] ){
        MM[jij]++;}
    if (jij<k-1 && !chain[jij] && !chain[jij+1] ){
        WW[jij]++;}
    }
    }

    /*
        //NOTE: Uncomment this section to view trains.
        //Also, it is worth reducing L to something small
//before doing so.
        trainPrint="";
        for (int iij=0; iij<k-1; iij++){
            trainPrint=trainPrint+ (chain[iij]?"0":"-");
        }
        System.out.println(trainPrint);
    */

    numsample++;
    if (numsample/1000==printseperator){
        // NOTE: this is where the output goes.
        //Output is seperated further and further apart as time
//goes on. This part of the function is frequently
        // rewritten, depending on what output is needed.
```

A.2. SimpleChainTest

```
        printseparator*=2;
// System.out.println("Train density"+ (float)(numTrainsLeave)/numsample
//        +" weight average"+TotalTotalWeight/numsample);

        double OldPredict=0;
        double cunningPredict=Mcount[0]/numsample;
        System.out.println("t"+ numsample);
        for (int iij=1; iij<k-1; iij++){
            //System.out.print("\n Height" + iij+" ");
            OldPredict=(Math.pow(r,iij)/M)/((Math.pow(r,iij)/M)+1-1/M);
            System.out.println(""+iij+ " & " + OldPredict+ " & "
                + Mcount[iij]/numsample+ " & " + (MM[iij]/Mcount[iij])
                    +"      \\\\" );

            /* //When uncommented this section
            //prints information about trains.
            //This was used to test the predictions made in
            // the paper against actual simulated results.

for (int ddd=1; ddd<k-1; ddd++){
    if (TrainArrive[iij][ddd]>0){
        System.out.print("length" + ddd+" occurance"
            +(double)(TrainArrive[iij][ddd])/numTrainsLeave
                + "Exptime"
            +(double)TrainTime[iij][ddd]/(TrainArrive[iij][ddd]) );
        }
    }*/
}

for (int iij=1; iij<k-1; iij++){
    OldPredict=(Math.pow(r,iij)/M)/((Math.pow(r,iij)/M)+1-1/M);
    cunningPredict=cunningPredict*(MM[iij]/Mcount[iij])
        + r*cunningPredict*(1-MM[iij]/Mcount[iij]);
    System.out.println(""+iij+ ";" + OldPredict+ ";" +
        Mcount[iij]/numsample+ ";" + (MM[iij]/Mcount[iij])
            + ";" +cunningPredict);
}
}
}
```

A.2. *SimpleChainTest*

}

}