Genetic diversity through reproductive semi-self-incompatibility

Bachelor's Project Mathematics

June 2024

Student: O.I. Daalder

First supervisor: Dr. S.A. Donderwinkel

Second assessor: Prof. dr. J.P. Trapman

Abstract

This thesis is discussing a model inspired by the 2021 paper by Smith et al. In that paper, the introduction of a self-incompatible mating type into a unisexual population is investigated. This thesis delves deeper into the dynamics of a population of semi-self-incompatible individuals: a transition state from unisexually reproducing individuals to completely self-incompatible individuals.

Contents

1 Introduction

One of the biggest challenges of sexually reproducing organisms is finding a suitable mate. Locating a mate is sometimes already a hard task, so finding out that you are of incompatible sex is a hindrance in the pursuit of reproducing. Biologically, it would therefore, naively, make sense that traits for indiscriminately mating are selected for, while individuals that have trouble finding mates are selected against. However, it is found that most organisms are not free to mate indiscriminately. In humans we see a two sexes system, a female and a male can produce offspring, while a male and male or a female and female are not able to do this. The sex chromosomes of *Animalia* are therefore what you would call *self-incompatible*.

In the 2021 paper by Smith et al. the case of a totally *unisexual* population is considered, a population where every individual is free to mate indiscriminately. In that paper, the mating type of individuals are discussed, which the reader can crudely think of as the ancestor of sex chromosomes. Most models of early evolution of sexual reproduction consider this unisexual mating type as the ancestor of the self-incompatible mating types [1]. The paper then investigates the introduction of self-incompatible (SI) mating types into this population and looks at which mating type sweeps into fixation with which parameters.

However, the model by Smith et al. gives rise to some questions. For example, the sudden introduction of SI mating types does not have a sound biological explanation. In an evolutionary sense, the self-incompatible mating type would make sense as slowly evolving from the unisexual type. A certain individual has lower chance of mating with its own type than mating with the rest of the types. This intermediate form could then give rise to a totally self-incompatible mating type. The goal of this thesis is to look at a *semi-self-incompatible* (SSI) mating type population and see what the dynamics of such a population are. This could give an intuition for further papers where from a unisexual population, gradually, self-incompatible mating types arise.

As this thesis views the problem through a mathematical lens, the use of biological jargon will be limited. Some terminology will be explained in this section, for the general understanding of the reader. A *haploid* population is a population where all the individuals have cells with a single set of chromosomes, as opposed to a *diploid* population, where cells have two sets of

chromosomes. In this thesis, the population that is discussed will always be of the haploid type. A DNA sequence is composed of *alleles*, which are different forms of a gene. Alleles can for example change eye colour, or influence nothing. Alleles are located at a place in the DNA sequence, called the *locus*. *Demes* refers to a certain subpopulation that can interbreed. When this thesis uses the term deme, you can think of the designated group being on another island than the main population, where interaction with other islands can occur through 'migration'. When considering parents in a population model, one parent is considered the 'searching' parent, that is called the *focal parent*.

A mathematical preliminary rooted in biology is the Moran model. First introduced in 1958 by Patrick Moran [2], this model is used to describe the stochastic behaviours of a population over time. In a neutral Moran model for a haploid population of size N, the total rate of reproduction events occurring is $\binom{N}{2}$. When such an event occurs, a pair of individuals is chosen uniformly at random from the population, where one individual dies, and the other produces one offspring. The offspring inherits the type of the parent and the population size stays the same.

The Moran model can be modified to model a sexually reproducing population. As a reproduction event occurs, a trio of individuals is chosen uniformly at random from the population, where one is chosen to die, and the offspring inherits genetic material from the two other individuals. This inheritance of genes is further explained in Section 2.

In this thesis, we will first look at the model introduced in the 2021 paper by Smith et al. and explain the slight changes we made to it in Section 2. After this, to aid the reader in the understanding of the rest of the paper, Section 3 explains diffusion processes as these are widely used in the field of population dynamics. In Section 4 we will state the main results found in this thesis and shortly review what they mean for the model. In the following section, Section 5, we will show how we achieved the main results. Finally, we will conclude and discuss on our findings in Section 6.

2 Model

In our model, we consider a population that can reproduce both asexually and sexually. This is therefore considered as a special case of the Moran

model, where if asexual reproduction occurs the model reverts back to the neutral Moran model, and otherwise it becomes the model briefly explained at the end of the introduction. For simplicity's sake, the rate at which an individual enters a reproductive phase is set to t. In our model, the ratio of sexual and asexual reproduction can be captured by $\bar{\nu} \in [0, 1]$, which will govern the proportion of the mating events that are sexual.

To be precise, an individual attempts to reproduce asexually at rate $t(1-\bar{\nu})$. Upon being chosen for an asexual reproduction event, a new individual is born that inherits the mating type of its parent. On the other hand, an individual reproduces sexually at rate $t\bar{\nu}$, where it inherits genetic material from both parents when the mating attempt is successful. The dynamics of this inheritance is explained in Section 2.3.

2.1 Replacement

As has been explained about the Moran model, in both type of reproduction events, the new individual replaces another individual in the population. A uniformly random individual is chosen to die, so that the new individual can take its place and the total amount of individuals stays equal to N.

2.2 Mating rounds

After having briefly introduced the different mating types, a clear distinction will be made between different mating types. Considering sexual reproduction, the self-incompatible (SI) mating type is a mating type that can mate with any mating type other than itself. For example, a mating type a individual can mate with individuals of type b, c, d, etc. but not with a type a individual. The unisexual mating type can mate indiscriminately, and therefore a sexual reproduction event always ends in success. The semi-self-incompatible (SSI) mating type is a hybrid of the previous two types, that can mate with any type, but has a preference for types other than its own.

During each sexual mating attempt, a potential mate is chosen uniformly at random from the $N - 1$ remaining individuals. In the original model by Smith et al. [3], the mating attempt of an SI individual would be rejected once it is instigated with an individual of the same mating type. In this thesis, that condition is relaxed for the SSI individuals. A parameter $\rho \in [0, 1]$ describes the probability that an attempt of an SSI individual reproducing with the same mating type is successful.

The number of attempts an individual gets at sexual reproduction is a random value A. Fixing a value $c > 0$, we choose A to be geometrically distributed with parameter $1 - e^{-c}$ leading to

$$\mathbb{P}[A > n] = e^{-cn}, \qquad n \geq 1.$$

This can be understood as, after each attempt, the individual decides to stop mating with probability $1 - e^{-c}$ independently. Therefore, the probability of the number of mating attempts being equal to n becomes

$$\mathbb{P}[A = n] = e^{-c(n-1)}(1 - e^{-c}), \qquad n \geq 1. \tag{2.1}$$

Note here that this probability stands for the individual giving up for the first time after the nth attempt.

Limits give interesting properties of this distribution. The limit $c \to \infty$ corresponds to a high mating cost for individuals, which results in them, asymptotically, only having one attempt, since $\mathbb{P}[A = 1]$ goes to one as $1 - e^{-c}$ grows close to one. The limit $c \to 0$ is the case where mating cost is low, and therefore the number of mating attempts go to infinity.

2.3   Recombination

When a sexual mating attempts ends in success, we need to understand which genes are inherited from which parent so that we know the effect of the reproduction event on the genetic composition of the population. As we consider a haploid population with distinct mating types, the mating type of an individual is determined by an allele at a single locus. The genome of the individuals in the population is considered to be so long, that we approximate the position of the mating type locus by a single point m within the normalized genome length $[0, 1]$.
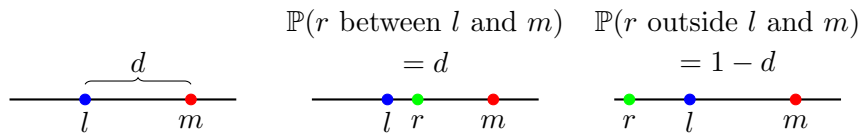

Figure 1: A quick visual representation of the inheritance of genes, if r is placed uniformly on the genome.

Progeny inherits genetic material from both parents as follows: a recombination site r is uniformly selected at random on the normalized genome. The offspring inherits the genetic material left of the chosen recombination point r from one random parent and the genetic material right of r from the other parent. In this thesis, a neutral genetic marker at position $l \in [0,1]$ on the genome is considered. We use this to track the effect of mating types on genetic diversity. The allelic type of this marker has no influence on an individual's mating type. The distance on the genome between the locus of the mating type and the locus of the neutral marker will be denoted by $d = |m - l|$. Therefore, the probability that mating type and the neutral marker are inherited by different parents is set to d, as the recombination site is chosen uniformly and the locus of the mating type m and neutral marker l are a distance d apart. When considering genetic diversity on a separate chromosome from the mating type, d is set to $\frac{1}{2}$. This ensures that there is no linkage between the neutral marker and mating type.
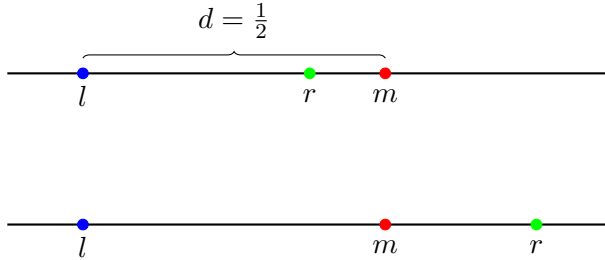

Figure 2: With d being equal to $\frac{1}{2}$, we can see that there is no linkage between neutral marker and mating type.

2.4 Mutation

Lastly, in correspondence to biological processes in nature, a mutation event will be added to the model, regardless if the reproduction is sexual or asexual. For simplicity's sake, the possibility of mutating mating types is not considered in this thesis and only mutation at the neutral locus is possible. If mutation occurs at this locus l, it mutates to a new allele not present in the population at that time (the neutral locus has choice of an endless set of alleles $\{a_1, a_2, ...\}$). For the probability of a mutation event occurring, the parameter $\bar{\mu} \in [0,1]$ is chosen. After any reproduction event (sexual or asexual) the genetic material is incorrectly transmitted (mutation) with

probability $\bar{\mu}$. Where $\bar{\mu} = 0$ would therefore be the artificial case that no mutation occurs in the model.

3   Diffusion process

To aid the reader in the understanding of the model, this section will introduce diffusion processes, as they play a big role in the field of genetics to describe the limiting evolution of a population as its size tends to infinity.

Definition 3.1 (Diffusion process). *A k-dimensional diffusion process is a Markov process in $\mathbb{R}^k$ with continuous sample paths. A diffusion process is described by the drift and diffusion coefficients, that describe the mean and covariance in the change of each coordinate $X_i(t)$ over an infinitesimal time period.*

A special case: a standard Brownian motion $(B_t^1, ..., B_t^k)$ multiplied by $\Sigma \in \mathbb{R}^{k \times k}$ and added to A for $A \in \mathbb{R}^k$ is a diffusion process with constant drift A and constant diffusion Σ.

By the Markov property, the drift and diffusion at time t can depend on the position of the diffusion at time t, but not on t. In population dynamics this makes sense, as the population's behaviour is influenced by the amount of individuals of type i there were at time t. Pick $i, j = 1, ..., k$ and $h > 0$, then set $\Delta_h X_i(t) := X_i(t + h) - X_i(t)$, for the change of X over the time interval $(t, t + h)$. The following equations, as seen in Chapter 1.3 of the lecture notes by Etheridge [4], then determine the elements of the drift and diffusion coefficients respectively.

Definition 3.2 (Drift and diffusion). *The drift coefficients for a diffusion process[1]:*

$$a_i(x) := \lim_{h \searrow 0} \frac{1}{h} \mathbb{E}[\Delta_h X_i(0) | X(0) = x]. \tag{3.1}$$

The diffusion coefficients for a diffusion process:

$$b_{ij}(x) := \lim_{h \searrow 0} \frac{1}{h} \mathbb{E}[\Delta_h X_i(0) \Delta_h X_j(0) | X(0) = x]. \tag{3.2}$$

---

[1]Note here that the a and b are swapped from the notation by Smith et al. This choice was made by the author as there was no apparent reason why b_i and a_{ij} were chosen by Smith et al. and a comes earlier in the alphabet than b

By looking at Definition 3.2, the following result is achieved about $b_{ij}(x)$.

Proposition 1. *The matrix $\mathbf{b}(x)$ is positive semi-definite for all x.*

Proof. If we take σ as the decomposition matrix of $\mathbf{b}$ and z_i as factors of the vector $\bar{z}$, we can say that

$$\sum_{i,j} z_i b_{ij} z_j = \sum_{i,j,k} z_i \sigma_{i,k} \sigma_{j,k} z_j = \sum_k \left(\sum_i z_i \sigma_{i,k} \right)^2 \geq 0.$$

$\square$

So $\mathbf{b}(x)$ is a positive semi-definitive matrix, which makes sense as it expresses the covariance in the change of each coordinate $X_i(t)$, and it is thus possible to be expressed as $\mathbf{b}(x) = \sigma(x)\sigma(x)^T$.

Then $\{X(t)\}_{t \geq 0}$ describes the solution to the system of stochastic differential equations

$$dX_i(t) = a_i(X(t))dt + \sum_{j=1}^k \sigma_{ij}(X(t))dB_j(t), \tag{3.3}$$

where $B_1, ..., B_k$ are independent Brownian motions. This can be approximated, for very small h, by

$$X_i(t+h) \approx X_i(t) + a_i(X(t))h + \sum_{j=1}^k \sigma_{ij}(X(t))\sqrt{h}\xi_j(t),$$

where, for each $j = 1, ..., k$ and $t \geq 0$, $\xi_j(t)$ is a standard normal random variable and $\xi_i(t)$ and $\xi_j(s)$ are independent if either $i \neq j$ or $t \neq s$.

Diffusion processes can occur as the limit of a sequence of Markov processes with increasingly frequent but diminishing jumps. This can be viewed as a generalisation of Donsker's Theorem, Theorem 8.1.4 in Probability theories and examples by Durrett [5], that states that scaling the time steps of a random walk by $\frac{1}{n}$ and the increment by $\frac{1}{\sqrt{n}}$ results in a Brownian motion. Consider a sequence of continuous time Markov processes $\{X^N\}_{N=1}^\infty$, where X^N takes values in $S_N \subseteq \mathbb{R}^k$. Then the following coefficients can be defined, which are the counterparts of drift and diffusion coefficients for a finite population.

$$a_i^N(x) := \lim_{h \searrow 0} \frac{1}{h} \mathbb{E}[\Delta_h X_i^N(0) | X^N(0) = x]$$

$$b_{ij}^N(x) := \lim_{h \searrow 0} \frac{1}{h} \mathbb{E}[\Delta_h X_i^N(0) \Delta_h X_j^N(0) | X^N(0) = x].$$

Theorem 3.1. *If for all $R < \infty$, a_i^N and b_{ij}^N get close to a_i and b_{ij} respectively as N goes to infinity and a tightness condition holds, then $X^N \Rightarrow X$ in $\mathbb{R}^k$ as $N \to \infty$. Here, $\Rightarrow$ denotes weak convergence in the space of càdlàg paths equipped with the Skorokhod topology.*

Where càdlàg paths stand for paths that are continuous from the right and converging from the left, and the Skorokhod topology informally is a topology of almost uniform convergence with the allowance of jumps being made [6]. The proof of this Theorem is beyond the scope of this thesis, and therefore the reader is referred to Theorem 7.1 and Lemma 8.2 in Durrett [7].

When we express this mathematically, this leads to the following three conditions:

$$\lim_{N \to \infty} \sup_{|x| \leq R} |a_i^N(x) - a_i(x)| = 0, \qquad \sup_{x \in [0,1]^k} |a_i(x)| < \infty,$$
$$\lim_{N \to \infty} \sup_{|x| \leq R} |b_{ij}^N(x) - b_{ij}(x)| = 0, \qquad \sup_{x \in [0,1]^k} |b_{ij}(x)| < \infty, \tag{3.4}$$

and

$$\lim_{N \to \infty} \sup_{|x| \leq R} \lim_{t \searrow 0} \frac{1}{t} \mathbb{E}[|X^N(t) - X^N(0)|^4 | X^N(0) = x] = 0. \tag{3.5}$$

This expression has an equivalent expression explained in Section 1.3 by Etheridge [4], where a sequence of transition rates, defined as

$$\frac{d}{dt} \mathbb{P}[X_t^N \in A | X_0^N]|_{t=0} = Q^N(x, A),$$

needs to satisfy the following three conditions for all $R > 0$:

$$\int_{\mathbb{R}^k} (y_i - x_i)(y_j - x_j) Q^N(x, dy) = b_{ij}^N(x) \to b_{ij},$$

$$\int_{\mathbb{R}^k} (y_i - x_i) Q^N(x, dy) = a_i^N(x) \to a_i(x)$$

and

$$\int_{\mathbb{R}^k} |y - x|^4 Q^N(x, dy) = \gamma^N \to 0. \tag{3.6}$$

Eq. 3.5 or, equivalently, Eq. 3.6 guarantees that the discrete process does not make jumps of macroscopic scale, which is a necessary condition for a limit process to be continuous.

4 Results

Two main results have been achieved in this thesis. One regards the stability of the fixed points of the proportions of mating types in the population. The other regards the diffusion process that the system investigating the genetic diversity in the mating type i deme approximates as N goes to infinity.

Theorem 4.1. *The stability of the fixed point* $\mathbf{p}^*$ *of the system decreases as* $\bar{\nu}$ *and* c *decrease or* k *and* ρ *increase.*

What is interesting to see is that all these values stand for an increase in successful reproduction. As $\bar{\nu}$ decreases, the population tends to reproduce asexually, which decreases the chance of a failed mating attempt. As c decreases, we have seen the limit of the amount of mating attempts go to infinity, which means that with more mating attempts, the total chance of one being successful increases. With k increasing, there are more mating types in the population and therefore the chance of picking the same mating type and a mating attempt failing is smaller. Lastly, as ρ increases, the population tends to a unisexual mating type population, from which we have seen that mating attempts always end in success.

Theorem 4.2. *As* N *goes to infinity, the system modelling genetic diversity tends to the solution of the stochastic differential equation*

$$X_i(t) = x_i - \int_0^t \mu X_i(s)ds + \int_0^t \sum_{\substack{j=1 \\ j \neq i}}^k m_{ij}(p, c)(X_j(s) - X_i(s))ds$$
$$+ \int_0^t \sqrt{\frac{2}{p_i} X_i(s)(1 - X_i(s))} dB_i(s). \tag{4.1}$$

What is interesting to note is that this differential equation is almost the same as the one proposed by Smith et al. However the slight difference is

nested in the $m_{ij}{}^2$, the transition rates from one mating type to another. We will see in Section 5.4 that only the factor with m_{ij} originates from sexual reproduction in the SDE and it would therefore make sense for it to be the only one to change when we change the dynamics of sexual reproduction.

5 Analysis

In this section, the model introduced in Section 2 will be investigated to prove the main Theorems given in the Results section. If the reader has forgotten about the way the model works, we urge them to go back and revisit Section 2 to ensure that this section will make sense.

5.1 Asexual reproduction

Asexual reproduction, which happens at rate $t(1-\bar{\nu})$ for a given individual, occurs at rate $tN(1-\bar{\nu})$ in a group of N individuals. When such an event occurs, a mating type i individual is picked from the population proportional to its frequency, $\bar{p}_i = \frac{N_i}{N}$, and an individual of mating type j is assigned death with probability proportional to its frequency, $\bar{p}_j$. Thus, the probability per unit time of mating type i increasing by one, as mating type j decreases by one during asexual reproduction is

$$tN(1-\bar{\nu})\frac{N_i}{N}\frac{N_j}{N}. \tag{5.1}$$

5.2 Sexual reproduction

For investigating the dynamics of the probability of mating type i increasing by one and type j decreasing by one through sexual reproduction per unit time, we need some extra analysis. In Section 2.2 we discussed the probability of a mating attempt being successful, here we will find a mathematical expression for this probability. At a given mating attempt, an SSI individual of the mating type i attempts to mate with an individual of mating type j with probability proportional to the individuals with mating type j. If j is different from i, the attempt is always successful. If $j = i$ the attempt succeeds with probability ρ. Creating a sum where the mth attempt is a

---

[2]Note here that m_{ij} denotes the transition rate of SSI individuals in this paper and m_{ij}^* denotes that of SI individuals as opposed to m_{ij} denoting the transition rate of SI individuals in Smith et al.

successful mating attempt, we need the first $m - 1$ mating attempts to be with the same mating type and probability $(1 - \rho)$ to ensure failure, this yields the equation

$$
\begin{aligned}
\theta_{ij}^N(c, \rho) &:= \sum_{n=1}^{\infty} \mathbb{P}[A = n] \sum_{m=1}^{n} \left((1 - \rho) \frac{N_i - 1}{N - 1} \right)^{m-1} \frac{N_j}{N - 1} \\
&= (1 - e^{-c}) e^c \sum_{n=1}^{\infty} e^{-cn} \frac{1 - \left((1 - \rho) \frac{N_i - 1}{N - 1} \right)^n}{1 - (1 - \rho) \frac{N_i - 1}{N - 1}} \frac{N_j}{N - 1} \\
&= \frac{1}{1 - (1 - \rho) e^{-c \frac{N_i - 1}{N - 1}}} \frac{N_j}{N - 1},
\end{aligned}
\tag{5.2}
$$

where the expression for $\mathbb{P}[A = n]$ comes from Eq. 2.1.

The second case is the probability that a reproductive event of a type i SSI individual results in a sexual reproduction with another type i individual. This will keep the equation almost completely the same, only that the mth successful mating attempt is with a type i individual with probability ρ for it being successful. This garners:

$$
\begin{aligned}
\theta_{ii}^N(c, \rho) &:= \sum_{n=1}^{\infty} \mathbb{P}[A = n] \sum_{m=1}^{n} \left((1 - \rho) \frac{N_i - 1}{N - 1} \right)^{m-1} \rho \frac{N_i - 1}{N - 1} \\
&= \frac{\rho}{1 - (1 - \rho) e^{-c \frac{N_i - 1}{N - 1}}} \frac{N_i - 1}{N - 1}.
\end{aligned}
\tag{5.3}
$$

Then the probability of a type i SSI individual finding a compatible mate in a population with k distinct mating types equals the sum of θ_{ij} with j ranging from 1 to k:

$$
\begin{aligned}
\theta_i^N(c, \rho) &:= \sum_{j=1}^{k} \theta_{ij}^N(c, \rho) = \theta_{ii}^N(c, \rho) + \sum_{\substack{j=1 \\ j \neq i}}^{k} \theta_{ij}^N(c, \rho) \\
&= \frac{\rho}{1 - (1 - \rho) e^{-c \frac{N_i - 1}{N - 1}}} \frac{N_i - 1}{N - 1} + \frac{1}{1 - (1 - \rho) e^{-c \frac{N_i - 1}{N - 1}}} \left(1 - \frac{N_i - 1}{N - 1} \right) \\
&= \frac{1}{1 - (1 - \rho) e^{-c \frac{N_i - 1}{N - 1}}} \left(1 - (1 - \rho) \frac{N_i - 1}{N - 1} \right).
\end{aligned}
$$

$$
\tag{5.4}
$$

14

Sexual reproduction attempts occur in the population at total rate $tN\bar{\nu}$. Three cases are considered for creating a progeny of type i: the first is that the focal parent is of type i and mates with another type i individual, which creates type i individuals with probability 1. The second is where a focal parent of type i mates with another type and produces type i offspring with probability $\frac{1}{2}$. The last is that the focal parent is of type $l \neq i$ and produces type i offspring with probability $\frac{1}{2}$ by mating with a type i individual. The equation expressing the gaining of a mating type i individual and the loss of a type j individual through one sexual reproductive event for k distinct mating types becomes

$$\frac{1}{2}tN\bar{\nu}\left[2\frac{N_i}{N}\theta_{ii}^N(c,\rho) + \sum_{l\neq i}^{k}\frac{N_i}{N}\theta_{il}^N(c,\rho) + \sum_{l\neq i}^{k}\frac{N_l}{N}\theta_{li}^N(c,\rho)\right]\frac{N_j}{N} \qquad (5.5)$$

with $\theta_{li}^N(c,\rho)$ from Eq. 5.2.

5.3 System dynamics and fixed points

For convenience, the reproductive rate t will be set to N. Summing Eq. 5.1 and 5.5, all cases of type i increasing and j decreasing through a reproductive event are expressed in one equation. Therefore the rate of the system transitioning from state $\mathbf{N} = (N_1, N_2, ..., N_{k-1})$ to $\mathbf{N}^{(i,j)} = (..., N_i+1, ..., N_j-1, ...)$ can be expressed as $T_{ij}(\mathbf{N})$, with

$$T_{ij}(\mathbf{N}) = N^2(1-\bar{\nu})\frac{N_i}{N}\frac{N_j}{N} + \frac{1}{2}N^2\bar{\nu}\left[2\frac{N_i}{N}\theta_{ii}^N(c,\rho) + \sum_{l\neq i}^{k}\frac{N_i}{N}\theta_{il}^N(c,\rho) + \sum_{l\neq i}^{k}\frac{N_l}{N}\theta_{li}^N(c,\rho)\right]\frac{N_j}{N}.$$

We want to study the probability of being in state $\mathbf{N}$ at time t. This is expressed as $\Phi_{\mathbf{N}}(t)$, which is given by equation

$$\frac{d\Phi_{\mathbf{N}}(t)}{dt} = \sum_{i,j}\left[T_{ji}(\mathbf{N}^{(i,j)})\Phi_{\mathbf{N}^{(i,j)}}(t) - T_{ij}(\mathbf{N})\Phi_{\mathbf{N}}(t)\right], \qquad (5.6)$$

where the net change is equal to the rate at which state $\mathbf{N}$ is entered minus the rate at which state $\mathbf{N}$ is left [8].

Now we consider a change of variables by dividing $\mathbf{N}$ by the number of individuals N, where $\bar{p}_i = \frac{N_i}{N}$ is the proportion of a certain mating type

in the population. This creates the states $\mathbf{p} = (\bar{p}_1, \bar{p}_2, ..., \bar{p}_{k-1})$ and $\mathbf{p}' = (..., \bar{p}_i + \frac{1}{N}, ..., \bar{p}_i - \frac{1}{N}, ...)$. Additionally, we make the transformation of $\tau = \frac{t}{N}$ [3] to 'slow' time so that macroscopic changes to the make-up of the population happen in $\Theta(1)$ units of time. Taking the Taylor expansion around $\frac{1}{N}$ and truncating at order $\mathcal{O}(N^{-1})$ (which is justified as we let $N \to \infty$), we find that the limiting density ϕ is described by:

$$\frac{\partial \Phi(\mathbf{p}, t)}{\partial \tau} = -\sum_{i=1}^{k-1} \frac{\partial}{\partial p_i}[A_i(\mathbf{p})\Phi(\mathbf{p}, \tau)]. \tag{5.7}$$

Here $A_i(\mathbf{p})$ is given by

$$
\begin{aligned}
A_i(\mathbf{p}) &= \lim_{N \to \infty} \sum_{j=1}^{k} \frac{1}{N^2}[T_{ij}(\mathbf{p}N) - T_{ji}(\mathbf{p}N)] \\
&= \frac{\bar{\nu}}{2} \lim_{N \to \infty} \sum_{j=1}^{k} \left[2\frac{N_i}{N}\frac{N_j}{N}\theta_{ii}^N(c, \rho) + \sum_{l \neq i}^{k} \frac{N_i}{N}\frac{N_j}{N}\theta_{il}^N(c, \rho) + \sum_{l \neq i}^{k} \frac{N_l}{N}\frac{N_j}{N}\theta_{li}^N(c, \rho) \right. \\
&\qquad \left. - 2\frac{N_j}{N}\frac{N_i}{N}\theta_{jj}^N(c, \rho) + \sum_{l \neq j}^{k} \frac{N_j}{N}\frac{N_i}{N}\theta_{jl}^N(c, \rho) + \sum_{l \neq j}^{k} \frac{N_l}{N}\frac{N_j}{N}\theta_{lj}^N(c, \rho) \right] \\
&= \frac{\bar{\nu}}{2} \sum_{j=1}^{k} \left[p_i p_j \left(\frac{1 - 2(\rho - 1)p_i}{1 - (1 - \rho)e^{-c}p_i} - \frac{1 - 2(\rho - 1)p_j}{1 - (1 - \rho)e^{-c}p_j} \right) \right]\Big|_{p_k = 1 - \sum_{j=1}^{k-1} p_j},
\end{aligned}
$$

where $p_i = \lim_{N \to \infty} \bar{p}_i$.

Eq. 5.7 is equivalent to the ODE

$$\frac{d\mathbf{p}}{d\tau} = A_i(\mathbf{p}). \tag{5.8}$$

From this system, we can find the fixed points for the proportions of the population. First, let us look at some results from linear algebra regarding fixed points of a system.

Definition 5.1 (Fixed point). *A fixed point of an ODE has derivatives w.r.t. time equal to zero.*

---

[3]This is a different notation from the paper by Smith et al. , where it is denoted by $\tau = tN$. This must have been a mistake in notation by Smith et al. as their calculations also use $\tau = \frac{t}{N}$.

Essentially, what this means for us is that we are looking for the answers to the equation:

$$A_i(\mathbf{p}) = \frac{\bar{\nu}}{2} \sum_{j=1}^{k} \left[p_i p_j \left(\frac{1 - 2(\rho - 1)p_i}{1 - (1 - \rho)e^{-c}p_i} - \frac{1 - 2(\rho - 1)p_j}{1 - (1 - \rho)e^{-c}p_j} \right) \right] \Big|_{p_k = 1 - \sum_{j=1}^{k-1} p_j}$$
$$= 0.$$
(5.9)

From this equation, we can derive that anything that picks $p_i = p_j$ will result in $A_i(\mathbf{p})$ being equal to zero. As p_i's stand for the proportions in the population, we need the sum of all these to be equal to 1. This leaves one option for the fixed point, where all p_i's are equal to $\frac{1}{k}$ for a population with k mating types. This would mean that each mating type is equally represented in the population. Interesting to check is whether this fixpoint is attracting (asymptotically stable), as this would mean the system will tend towards this point.

Theorem 5.1. *A fixed point of a system of differential equations is asymptotically stable if the real parts of the eigenvalues of the Jacobian of the system evaluated at the fixed point are negative.*

The proof of this Theorem is beyond the scope of this paper, but it can be found in Chapter 3.1 in Differential Equations, Dynamical Systems, and an Introduction to Chaos, Second Edition by R. Devaney and M.W. Hirsch[9]. When calculating the Jacobian at the fixed point $\mathbf{p}^*$, we get a $(k-1) \times (k-1)$ matrix looking like

$$J_{ii} = \frac{\bar{\nu}2e^{2c}(2e^c + 1)(\rho - 1)}{((\rho - 1)\frac{1}{k} + e^c)^3} \ \forall \ i,$$
$$J_{ij} = 0 \ \forall \ i \neq j.$$
(5.10)

Note here that the Jacobian is already diagonal, so the eigenvalues are the entries on the diagonal. The denominator of the fraction is positive and $\bar{\nu}2e^{2c}(2e^c + 1)$ also is, and since $\rho \in [0, 1]$ it means that $(\rho - 1)$ is negative and therefore the eigenvalues of the system are non-negative. Note here that the case where the eigenvalues are zero is the case that ρ is equal to 1 and we are dealing with a unisexual population. This would mean that it is an attractive fixed point, and the system will tend to it if given the chance.

Corollary 5.1.1. *The stability of a fixed point decreases as the real parts of the eigenvalues of the Jacobian evaluated at the fixed point get bigger (closer to zero).*

With this lemma, the proof of Theorem 4.1 is complete. From Eq. 5.10 we can deduce that the stability decreases as $\bar{\nu}$ and c decrease and k and ρ increase.

As this paper looks into the dynamics of SSI mating types, it is interesting to say something about the results regarding ρ. As ρ increases and the eigenvalues get closer to zero, the population tends more to a unisexually reproducing population. In a totally unisexual population mating types are considered obsolete, as anyone can mate with anyone regardless of what mating type they have. Since mating attempts are always successful in this system, we would expect more erratic behaviour in the proportions of the population, and $\frac{N}{k}$ would not represent a fixed point.

When a mutation event occurs that introduces a new type ($k \to k + 1$), the fixed point (with $\bar{p}_{k+1} \approx 0$) becomes unstable, however the fixed point with $\bar{p}_i = \frac{1}{k+1}$ for all i is stable with eigenvalues equal to those in Eq. 5.10, only then with $k + 1$ instead of k.

5.4 Diffusion approximation

From now on, it is assumed that both mutation at locus l and sex are rare, as this assumption makes it able to find interesting results about the system. It will be assumed that for $\nu > 0$ and $\mu > 0$,

$$\bar{\nu} = \frac{\nu}{N},$$

and

$$\bar{\mu} = \frac{\mu}{N},$$

where ν is interpreted as the average number of sexual mating rounds per N reproductive events, and μ is the average number of successful reproductive events with mutation per N successful reproductive events.

Recall that we have introduced a neutral genetic marker at locus l. This marker can contain one allele of an infinite set of alleles $\{a_1, a_2, ...\}$. To track the genetic diversity in the mating type i deme, we define the following equation.

Definition 5.2.

$$X_i^N(t) := \frac{\#\{\textit{Type i individuals carrying allele } a_1 \textit{ at locus l at time t}\}}{N_i(t)},$$

(5.11)

is a Markov process that tracks the proportion of individuals of the ith mating type carrying the allele a_1 at the neutral marker at time t.

As we have seen in Section 5.4, $X_i^N(t)$ will approximate a diffusion process. This means that it will converge to some $X_i(t)$, which is the solution of the stochastic differential equation introduced in Theorem 4.2.

In Eq. 5.11 we can see that both the numerator and the denominator will change over time in the system. In order to make only one of these depend on time, we will fix the proportion of type i individuals. By using our knowledge from Section 5.3, we can set $N_i(t)$ to $\frac{N}{k}$, as we know that the fixed point for proportions $\frac{1}{k}$ is attracting. By making the population size N go to infinity, relatively little information is lost, and this seems a good assumption to simplify our model.

From Definition 5.2, certain terms arise for $a_i^N(x)$ and $b_{ij}^N(x)$ with $x = (x_1, ..., x_k)$. As their derivation is very similar, the derivation of $a_i^N(x)$ will be highlighted and then some changes will be explained to arrive at $b_{ij}^N(x)$.

The terms of $a_i^N(x)$ stand for each possible outcome of an individual's reproductive phase. Let us first consider the case of asexual reproduction. Here, individuals of mating type i carrying a_1 at the neutral locus l can be gained by asexual reproduction without mutation. They can also be lost by being replaced by mutation, or by being chosen to die as an individual of another type reproduces asexually (in which case the mutation of this other individual does not matter, as it will replace the i type individual nonetheless). This yields the following expression for a first term of $a_i^N(x)$, with x_i denoting the initial frequency of a_1 alleles in the ith mating type deme:

$$\frac{1}{N_i}N_i x_i N(1-\bar{\nu})(1-\bar{\mu})(1-x_i) - \frac{1}{N_i}N_i x_i N(1-\bar{\nu})\bar{\mu}x_i - \frac{1}{N_i}N_i(1-x_i)N(1-\bar{\nu})x_i,$$

which simplifies to

$$-\left(1 - \frac{\nu}{N}\right)\mu x_i.$$

The following six cases are the cases where individuals of type i with allele a_1 at l increase or decrease by $\frac{1}{N_i}$ through sexual reproduction without mutation. For clarity, we will refer to individuals of type i with allele a_1 at l as *tracked* individuals from now on. The first three cases are those where a tracked individual is gained. So there is no mutation, a non-tracked individual is picked to be replaced (with probability $1 - x_i$) and the probability of inheriting type i conditional on mating with a type j individual is given by $\frac{1}{2}(dx_j + (1 - d)x_i)$. One case considers a type i individual as the focal parent, another considers type j the focal parent, and the third considers a type i individual mating with another individual of type i:

$$\frac{1}{N_i}N\bar{\nu}N_i\sum_{\substack{j=1\\j\neq i}}^{k}\theta_{ij}^N(c,\rho)\tfrac{1}{2}(dx_j + (1-d)x_i)(1-\bar{\mu})(1-x_i)$$

$$+\frac{1}{N_i}N\bar{\nu}\sum_{\substack{j=1\\j\neq i}}^{k}N_j\theta_{ji}^N(c,\rho)\tfrac{1}{2}(dx_j + (1-d)x_i)(1-\bar{\mu})(1-x_i).$$

$$+\frac{1}{N_i}N\bar{\nu}N_i\theta_{ii}^N(c,\rho)x_i(1-\bar{\mu})(1-x_i)$$

The other three cases are those where a tracked individual is lost, so the typing or allele is inherited from the non-tracked individual and a tracked individual is assigned death, resulting in:

$$-\frac{1}{N_i}N\bar{\nu}N_i\sum_{\substack{j=1\\j\neq i}}^{k}\theta_{ij}^N(c,\rho)\tfrac{1}{2}(d(1-x_j)+(1-d)(1-x_i))(1-\bar{\mu})x_i$$

$$-\frac{1}{N_i}N\bar{\nu}\sum_{\substack{j=1\\j\neq i}}^{k}N_j\theta_{ji}^N(c,\rho)\tfrac{1}{2}(d(1-x_j)+(1-d)(1-x_i))(1-\bar{\mu})x_i$$

$$-\frac{1}{N_i}N\bar{\nu}\theta_{ii}^N(c,\rho)(1-x_i)(1-\bar{\mu})x_i$$

These six terms sum to

$$\left(1-\frac{\mu}{N}\right)\frac{\nu d}{2}\left[\sum_{j\neq i}^{k}\frac{N_i\theta_{ij}^N(c,\rho)+N_j\theta_{ji}^N(c,\rho)}{N_i}\right](x_j - x_i).$$

Then the last three terms cover the cases where during a sexual reproduction phase, a tracked individual is lost due to mutation. One where the focal parent is of type i and the other of type j and the last where a type i individual mates with another type i individual.

$$-\frac{1}{N_i} N\bar{\nu} N_i \sum_{\substack{j=1 \\ j\neq i}}^{k} \theta_{ij}^N(c,\rho) \frac{1}{2}\bar{\mu} x_i$$

$$-\frac{1}{N_i} N\bar{\nu} \sum_{\substack{j=1 \\ j\neq i}}^{k} N_j \theta_{ji}^N(c,\rho) \frac{1}{2}\bar{\mu} x_i$$

$$-\frac{1}{N_i} N\bar{\nu} N_i \theta_{ii}^N(c,\rho) \bar{\mu} x_i$$

This simplifies to

$$-\frac{\mu}{N}\frac{\nu}{2} x_i \left[2\theta_{ii}^N(c,\rho) + \sum_{j\neq i}^{k} \frac{N_i \theta_{ij}^N(c,\rho) + N_j \theta_{ji}^N(c,\rho)}{N_i} \right].$$

Proposition 2. *The counterpart of the drift coefficient for finite populations:*

$$a_i^N(x) = -\left(1-\frac{\nu}{N}\right)\mu x_i + \left(1-\frac{\mu}{N}\right)\frac{\nu d}{2} \left[\sum_{j\neq i}^{k} \frac{N_i \theta_{ij}^N(c,\rho) + N_j \theta_{ji}^N(c,\rho)}{N_i} \right](x_j - x_i)$$

$$-\frac{\mu}{N}\frac{\nu}{2} x_i \left[2\theta_{ii}^N(c,\rho) + \sum_{j\neq i}^{k} \frac{N_i \theta_{ij}^N(c,\rho) + N_j \theta_{ji}^N(c,\rho)}{N_i} \right].$$

Proposition 3. *When N goes to infinity, the ith factor of the drift coefficient of $X(t)$ is expressed as*

$$a_i(x) := -\mu x_i + \frac{\nu d}{2} \sum_{j\neq i}^{k} \frac{p_i \theta_{ij}(c,\rho) + p_j \theta_{ji}(c,\rho)}{p_i}(x_j - x_i), \qquad (5.12)$$

where $\theta_{ij}(c,\rho)$ and $\theta_{ii}(c,\rho)$ are defined by letting N go to infinity in Eq. 5.2 and Eq. 5.3, resulting in:

$$\theta_{ij}(c,\rho) = \frac{1}{1-(1-\rho)p_i e^{-c}} p_j \qquad (5.13)$$

$$\theta_{ii}(c,\rho) = \frac{\rho}{1-(1-\rho)p_i e^{-c}} p_i.$$

Proof. To check that $a_i^N(x)$ converges to this equation, it is needed to check condition 3.4.

Firstly, let us check that $\sup_{x\in[0,1]^k} |a_i(x)| < \infty$. This boundedness follows because variables μ, x_i, ν and d in the sum stay between 0 and 1. Now check the second condition:

$$|a_i^N(x) - a_i(x)| \le \frac{\nu}{N}\mu x_i$$
$$+ \frac{\nu d}{2}\left[\sum_{j\neq i}^{k}\left|\frac{N_i\theta_{ij}^N(c,\rho) + N_j\theta_{ij}^N(c,\rho)}{N_i} - \frac{p_i\theta_{ij}(c,\rho) + p_j\theta_{ij}(c,\rho)}{p_i}\right|\right]|x_j - x_i|$$
$$+ \frac{\mu}{N}\nu\theta_{ii}^N x_i$$
$$+ \frac{\mu}{N}\frac{\nu}{2}\left[\sum_{j\neq i}^{k}\frac{N_i\theta_{ij}^N(c,\rho) + N_j\theta_{ij}^N(c,\rho)}{N_i}\right](d|x_j - x_i| + x_i),$$

(5.14)

Which goes to zero because the first, third and fourth element go to zero as N goes to infinity as we divide by N. The second element goes to zero because as θ_{ij} is the limit of θ_{ij}^N as N goes to infinity. □

When looking at the $b_{ij}^N(x)$ case, only $b_{ii}^N(x)$ elements are considered, as $b_{ij}^N(x) = 0$ for all $i \neq j$ because we keep the demes at equal size ($\frac{N}{k}$). These are derived in the same way as before but now instead of plus or minus $\frac{1}{N_i}$ all first elements turn into $+\frac{1}{N_i^2}$. Then these new sums cancel to

$$2\left(1 - \frac{\nu}{N}\right)\left(1 - \frac{\mu}{N}\right)\frac{N}{N_i}x_i(1 - x_i) + \frac{\mu}{N}\left(1 - \frac{\nu}{N}\right)\frac{N}{N_i}x_i,$$

$$\frac{1}{N_i}\left(1 - \frac{\mu}{N}\right)\frac{\nu}{2}4\theta_{ii}^N(c,\rho)x_i(1 - x_i) +$$

$$\frac{1}{N_i}\left(1 - \frac{\mu}{N}\right)\frac{\nu}{2}\sum_{j\neq i}^{k}(2x_i(1 - x_i) + d(1 - 2x_i)(x_j - x_i))\frac{N_i\theta_{ij}^N(c,\rho) + N_j\theta_{ji}^N(c,\rho)}{N_i},$$

and

$$\frac{1}{N_i}\frac{\mu}{N}\frac{\nu}{2}x_i\left[2\theta_{ii}^N(c,\rho)+\sum_{j\neq i}^{k}\frac{N_i\theta_{ij}^N(c,\rho)+N_j\theta_{ji}^N(c,\rho)}{N_i}\right].$$

Proposition 4. *The counterpart of the diffusion coefficient for finite populations:*

$$b_{ii}^N(x):=2\left(1-\frac{\nu}{N}\right)\left(1-\frac{\mu}{N}\right)\frac{N}{N_i}x_i(1-x_i)+\frac{\mu}{N}\left(1-\frac{\nu}{N}\right)\frac{N}{N_i}x_i$$

$$\frac{1}{N_i}\left(1-\frac{\mu}{N}\right)\frac{\nu}{2}4\theta_{ii}^N(c,\rho)x_i(1-x_i)+$$

$$\frac{1}{N_i}\left(1-\frac{\mu}{N}\right)\frac{\nu}{2}\sum_{j\neq i}^{k}(2x_i(1-x_i)+d(1-2x_i)(x_j-x_i))\frac{N_i\theta_{ij}^N(c,\rho)+N_j\theta_{ji}^N(c,\rho)}{N_i}+$$

$$\frac{1}{N_i}\frac{\mu}{N}\frac{\nu}{2}x_i\left[2\theta_{ii}^N(c,\rho)+\sum_{j\neq i}^{k}\frac{N_i\theta_{ij}^N(c,\rho)+N_j\theta_{ji}^N(c,\rho)}{N_i}\right].$$

$$(5.15)$$

Proposition 5. *This means that as N goes to infinity, the equation for $b_{ij}(x)$ becomes:*

$$b_{ij}(x):=\frac{2}{p_i}x_i(1-x_i),\qquad j=i$$

$$0,\qquad j\neq i.$$

$$(5.16)$$

Proof. Once again, to check that $b_{ii}^N(x)$ indeed goes to $b_{ii}(x)$, check condition 3.4. Firstly, $\sup_{x\in[0,1]^k}|b_{ij}(x)|<\infty$ holds as all elements are between 0 and 1. Then,

$$|b_{ij}^N(x)-b_{ij}(x)|\leq$$

$$|2\left(1-\frac{\nu}{N}\right)\left(1-\frac{\mu}{N}\right)\frac{N}{N_i}x_i(1-x_i)+\frac{\mu}{N}\left(1-\frac{\nu}{N}\right)\frac{N}{N_i}x_i-\frac{2}{p_i}x_i(1-x_i)|$$

$$+\frac{1}{N_i}\frac{\nu}{2}\left[4x_i(1-x_i)\theta_{ii}^N(c,\rho)+\sum_{j\neq i}^{k}(2x_i(1-x_i)+d|1-2x_i||x_j-x_i|)\frac{N_i\theta_{ij}^N(c,\rho)+N_j\theta_{ij}^N(c,\rho)}{N_i}\right]$$

$$+\frac{1}{N_i}\frac{\mu}{N}\frac{\nu}{2}\left[(2x_i(1-x_i)+x_i)\theta_{ii}^N(c,\rho)\right.$$

$$+\sum_{j\neq i}^{k}(2x_i(1-x_i)+d|1-2x_i||x_j-x_i|+x_i)\frac{N_i\theta_{ij}^N(c,\rho)+N_j\theta_{ij}^N(c,\rho)}{N_i}\right]$$

This equation once again goes to zero as N goes to infinity, by the same logic as Eq. 5.14. □

Lastly, it is needed to check the tightness condition given in 3.5. Take γ^N as defined in 3.6. It follows from Jensen's inequality, Theorem 1.5.1 in Probability theory and examples by Durrett[5], that

$$|x - y|^4 \leq k \sum_{i=1}^{k} |x_i - y_i|^4,$$

therefore

$$\gamma^N(x) \leq k \sum_{i=1}^{k} \int_{\mathbb{R}^k} |x_i - y_i|^4 Q^N(x, dy) = k \sum_{i=1}^{k} \frac{1}{N_i^2} b_{ii}^N(x) = k^3 \sum_{i=1}^{k} \frac{1}{N^2} b_{ii}^N(x).$$

We can see that the right-hand side of this equation goes to zero, so $\gamma^N(x) = 0$ and the tightness condition is verified.

In this scenario, as $N \to \infty$, the allele frequencies within mating type are well approximated by a k-dimensional diffusion process which is a solution to a stochastic differential equation. From Eq. 5.12 and 5.16, the following SDE arises:

$$
\begin{aligned}
X_i(t) = x_i &- \int_0^t \mu X_i(s) ds + \int_0^t \sum_{\substack{j=1 \\ j \neq i}}^{k} m_{ij}(p, c)(X_j(s) - X_i(s)) ds \\
&+ \int_0^t \sqrt{\frac{2}{p_i} X_i(s)(1 - X_i(s))} dB_i(s),
\end{aligned}
\tag{5.17}
$$

where B_i's are independent Brownian motions and $m_{ij}(p, c)$ is the migration rate of from SSI mating type i to SSI mating type j. As seen in Eq. 5.12, $m_{ij}(p, c)$ is characterised by

$$
m_{ij}(p, c) :=
\begin{cases}
\dfrac{\nu d}{2} \dfrac{p_i \theta_{ij}(c, \rho) + p_j \theta_{ji}(c, \rho)}{p_i}, & i \neq j, \\
0, & j = i.
\end{cases}
$$

Which can be rewritten with Eq. 5.13 as:

$$m_{ij}(p,c) = \frac{\nu d}{2} \left(\frac{1}{1 - (1-\rho)e^{-c}p_i} + \frac{1}{1 - (1-\rho)e^{-c}p_j} \right) p_j, \qquad i \neq j.$$
(5.18)

This is slightly different from the SDE given in the paper by Smith et al. for SI mating types. The difference lies in the factor containing the migration rate, which makes sense biologically, as the other two factors arise from asexual reproduction and this is the same in both mating types. The migration rates for SI mating types look like:

$$m_{ij}^{*}(p,c) = \frac{\nu d}{2} \left(\frac{1}{1 - e^{-c}p_i} + \frac{1}{1 - e^{-c}p_j} \right) p_j, \qquad i \neq j.$$
(5.19)

Here we can see that if ρ is zero, Eq. 5.18 reverts to Eq. 5.19 as we are looking at the SI case. As ρ goes to one, Eq. 5.18 becomes $\nu d p_j$, which loses any factor including i and stops making sense as we are looking at a unisexual population where mating types are obsolete. With ρ anywhere between zero and one, we see that the transition rates become lower. This makes sense in the model, as mating types can now mate with the same mating type and therefore keep their typing. This lowers the probability of transitioning to another mating type.

6 Discussion & Conclusion

As we have seen from the Theorems in this thesis, interesting results have been found on the dynamics of a totally SSI population. We can conclude from Theorem 4.1 that the population prefers to go to unisexual reproduction if given the chance. This biologically makes sense, as this causes more successful reproduction in the population. We have also seen that the transition rates from mating type i to type j are influenced by semi-self-incompatibility. The probability of transitioning from one to another type becomes smaller, as there is an added probability of mating with the same type and not transitioning at all.

What is interesting to note is the minimal effect the introduction of semi-self-incompatibility appears to have on the genetic diversity. However, this can be explained by discussing an assumption we made in Section 5.4. In the beginning of that section, we explained that we keep the proportions of type i individuals equal to $\frac{N}{k}$. This causes the elements b_{ij} with $j \neq i$ in

Eq. 5.16 to be equal to zero, since we cannot lose a type j individual when a type i individual is reproducing, as we are forced to lose another type i individual with these assumptions. Although this choice can be justified with the population growing to infinity and the fixed point being at these proportions, we can still conclude that a lot of information is lost. In future research, we therefore urge to not make this assumption, to more accurately express the dynamics of the system.

As the extension of the model to include semi-self-incompatibility was quite simply, it gives rise to the possibility of other models. We can for example think of some mating types being less compatible with other mating types that are not their own. In this way, we can think of mating types as a spectrum, where interaction between mating types is based on how compatible they are. We can translate this to the following mathematical model. Uniformly assign every mating type a unique value from the unit interval $[0, 1]$. Then the compatibility of the individuals is determined by a function of either the distance of these numbers $g(z) = g(|x - y|)$ or the numbers $f(x, y)$. Obviously this function is a probability of mating attempt being successful and therefore maps to the unit interval. Initially, we can assume that these functions are smooth, to make mathematical results and conclusions follow easier. However, as natural processes are unpredictable, the functions may show jumps. With these functions, we can express the new model in the following way: instead of ρ, we now put the function for the probability of a mating attempt being successful in the model.

In conclusion, the field of stochastic differential equations is a fascinating mathematical field with a lot of real life applications. From telecommunications to finance to of course population dynamics. With this thesis, we have investigated the evolutionary case of semi-self-incompatible mating types, created as an enrichment of the 2021 model by Smith et al. As Smith et al. repeatedly state in their paper, there are several other models that investigate the early evolution of mating types in species. These other models make other assumptions about their populations and simplify their equations in other ways than Smith et al. do, as their angle of research is differently motivated [10]. What is good to keep in mind in future research is whether simplifications of the model are correctly made and do not cause any essential information to be lost or misinterpreted.

Acknowledgements

References

[1] Joseph Heitman. Evolution of sexual reproduction: a view from the fungal kingdom supports an evolutionary epoch with sex before sexes. *Fungal Biology Reviews*, 29(3-4):108–117, 2015.

[2] Patrick Alfred Pierce Moran. Random processes in genetics. In *Mathematical proceedings of the cambridge philosophical society*, volume 54, pages 60–71. Cambridge University Press, 1958.

[3] ASA Smith, S Penington, I Letter, DB Wilson, and GWA Constable. Genetic diversity in facultatively sexual populations and its implications for the origins of self-incompatibility in algae and fungi. *bioRxiv*, pages 2021–04, 2021.

[4] Alison Etheridge. Spatial population models. *Ensaios Matematicos*, 34, 2019.

[5] Rick Durrett. *Probability: theory and examples*, volume 49. Cambridge university press, 2019.

[6] Julian Kern. Skorokhod topologies: What they are and why we should care. *Mathematische Semesterberichte*, 71(1):1–18, 2024.

[7] Richard Durrett. *Stochastic calculus: a practical introduction*. CRC press, 2018.

[8] Alan J McKane, Tommaso Biancalani, and Tim Rogers. Stochastic pattern formation and spontaneous polarisation: the linear noise approximation and beyond. *Bulletin of mathematical biology*, 76:895–921, 2014.

[9] Morris W Hirsch, Stephen Smale, and Robert L Devaney. *Differential equations, dynamical systems, and an introduction to chaos*. Academic press, 2013.

[10] George WA Constable and Hanna Kokko. The rate of facultative sex governs the number of expected mating types in isogamous species. *Nature Ecology & Evolution*, 2(7):1168–1175, 2018.