A spatial mutation model with increasing mutation rates

Brian Chao

Advisor: Professor Jason Schweinsberg

May 31, 2021

Abstract

In this paper we consider a spatial model where cells are points on a d-dimensional torus $\mathcal{T} = [0, L]^d$; each unmutated cell incrementally mutates to types $k \in \mathbb{N}$ according to a collection of spatial Poisson processes Π_k with (mutation) rates μ_k. We will assume that μ_k are increasing, and find the asymptotic waiting time for the first mutation of type k to occur as the torus volume tends to infinity. This paper generalizes results on waiting for $k \geq 3$ mutations from the paper "Mutation Timing in a Spatial Model of Evolution" by Foo, Leder, and Schweinsberg.

1   Introduction

Cancer is caused by genetic mutations which disrupt regular cell division and apoptosis, in which case cancerous cells divide too rapidly and healthy cells reproduce too slowly. This can happen, for example, as soon as several distinct mutations occur and dramatically disrupt cell function. Thus, it is reasonable to model cancer as occurring after k distinct mutations appear in a sequence within a large body.

There have been mathematical models wherein cancer occurs once an individual first obtains a type $k \in \mathbb{N}$ cancerous cell after having already obtained cancerous cells of types $1, 2, ..., k - 1$. Such models date back to the celebrated 1954 paper by Armitage and Doll [1], which proposes a multi-stage model of carcinogenesis: an individual contacts mutations of types 1, 2, 3, ... in that order, and once an individual has a mutation of type $k - 1$ at time t, the probability of contacting a mutation of type k in a small time interval $(t, t + dt)$ is

$$\frac{\mu_1 \mu_2 \cdots \mu_k t^{k-1}}{(k-1)!} dt.$$

That is, the incidence rate of kth mutation (at which point the individual becomes cancerous) is proportional to $\mu_1 \mu_2 \cdots \mu_k t^{k-1}$. This means that cancer risk is positively correlated with both the mutation rates and the $(k - 1)$th power of age.

In 1972, Williams and Bjerknes [8] introduced the biased voter model as a model of skin cancer. To each site of epithelial tissue, there is an associated binary state representing if the site is cancerous or healthy. The model is biased in the sense that a cancerous cell spreads $\kappa > 1$ times as fast as a healthy cell; the constant κ is called the carcinogenic advantage. Cancer spread in the basal layer (i.e. hexagonal lattice), with exponential rate of spread, was computer-simulated.

1

More recently in 2016, Durrett, Foo, and Leder [4] worked on a spatial Moran model which is a generalization of the biased voter model. Cells are modelled as points of the integer lattice $(\mathbb{Z} \bmod L)^d$, and each cell is of type $i \in \mathbb{N} \cup \{0\}$. Higher cell types correspond to greater fitness. More specificially, type i cells have fitness level $(1 + s)^i$, where $s > 0$ measures the selective advantage of one cell over its predecessors. In the same paper, to simplify the analysis on finding the first time a second mutation occurs in $(\mathbb{Z} \bmod L)^d$, the authors also introduce a continuous model where cells live inside the torus $[0, L]^d$. The continuous stochastic model well-approximates the biased voter model because of the Bramson-Griffeath shape theorem [2, 3], which implies that the cluster of mutations in $\mathbb{Z}^d$ grows to the shape of a convex subset of $\mathbb{R}^d$. Whereas in the discrete model mutations only spread to adjacent sites in integer lattice, in the continuous model mutations spread uniformly in a ball, which simplifies the model. In Section 4 of [4], the authors use the continuous model to compute σ_2, the first time a type 2 mutation occurs, under certain asymptotic assumptions on the mutation rates. We use σ_k to denote the first time a type k mutation appears, after all mutations of types up to $k - 1$ have already appeared.

In 2020, Foo, Leder, and Schweinsberg [5] generalized the results in [4] by calculating the asymptotic distribution of σ_k for $k \geq 2$ assuming equal mutation rates $\mu_i = \mu$ for all i. The model used in [5] is essentially the model in [4]. Cancer spread is modeled on the d-dimensional torus $\mathcal{T} := [0, L]^d$ (continuous analog of $(\mathbb{Z} \bmod L)^d$ in [4]). We write $N := L^d$ to denote the volume of $\mathcal{T}$. Each point in $\mathcal{T}$ is assigned a mutation type. At initial time $t = 0$, all points in $\mathcal{T}$ are type 0, i.e. have no mutations. A type 1 mutation then occurs at rate μ_1 per unit volume. Once each type 1 mutation appears, it spreads out in a ball at rate α per unit time. This means that after the first type 1 mutation appears at time σ_1 and location $x \in \mathcal{T}$, all points in a ball of radius αt centered at x will also acquire a type 1 mutation by time $\sigma_1 + t$. Type 1 points then acquire a type 2 mutation at rate μ_2 per unit volume, and this process continues indefinitely. In general, type k mutations overtake type $k - 1$ mutations at rate μ_k per unit volume, and each type k mutation then grows out in a ball at rate α per unit time.

In this paper, we will further generalize the results in [5] in the case where the mutation rates are increasing, and either $\mu_1 \ll \alpha/N^{(d+1)/d}$ or $\mu_1 \gg \alpha/N^{(d+1)/d}$. In the first case, the mutations appear slowly enough so that σ_k is a sum of independent exponential random variables. In the second case, the mutations appear sufficiently quickly, and the waiting time σ_k depends only on if the rest of the mutations other than the type 1 mutation happen even more quickly.

On the other hand when $\mu_1 \asymp \alpha/N^{(d+1)/d}$, there will be many overlaps in between the type 1 regions, and we would have to compute the volume of regions consisting of many overlapping balls. Hence we do not pursue this case in our paper. Note that if $\mu_1 \asymp \alpha/N^{(d+1)/d}$ and all mutation rates are equal (i.e. $\mu_i = \mu$ for all i), it is proven, as a special case of Theorem 12 in [5], that $N\mu\sigma_k$ converges in distribution to a nondegenerate random variable for every $k \geq 1$.

As in [5], we will continue to assume that the rate of mutation spread α is constant across mutation types, so that successive mutations have equal selective advantage. One possible generalization of our model is by allowing each type i mutation to have a different rate of spread α_i, where $(\alpha_i)_{i=1}^{\infty}$ is decreasing. We could also allow $\alpha_{i+1} > \alpha_i$, but then regions of type $i + 1$ could quickly swallow the type i region. Subsequently, we have to study not only how the mutations of types $i + 1$ and i compete, but also how mutations of types $i + 1$ and $j \in \{1, ..., i - 1\}$ compete, making finding σ_k a more complex problem. We do not pursue this generalization.

We mention two biological justifications to assuming increasing mutation rates. Loeb and Loeb [6] suggest a general phenomenon in carcinogenesis where there is favorable selection for certain mutations, namely those that promote tumor growth in genes responsible for repairing

DNA damage. The increasing genetic instability disrupting DNA repair, in the context of this paper, would correspond to the effect of increasing mutation rates. Also our model in this paper would be of interest in the situation described in Prindle, Fox, and Loeb [7], where increasing mutation rates could actually incapacitate or kill malignant cells, which could then be applied toward improving chemotherapy methods.

In Section 2, we introduce basic notation, as well as a summary of results with heuristics for why the results should hold true. In Section 3, we introduce the space-time structure of the model and prove the limit theorems from Section 2.

2 Waiting for k mutations: results and heuristics

Given two sequences $(a_N)_{N=1}^\infty$ and $(b_N)_{N=1}^\infty$, we write:

1. $a_N \sim b_N$ if $\lim_{N\to\infty} a_N/b_N = 1$;

2. $a_N \ll b_N$ if $\lim_{N\to\infty} a_N/b_N = 0$ and $a_N \gg b_N$ if $\lim_{N\to\infty} a_N/b_N = \infty$;

3. $a_N \asymp b_N$ if $0 < \liminf_{N\to\infty} a_N/b_N \leq \limsup_{N\to\infty} a_N/b_N < \infty$;

4. $a_N \lesssim b_N$ if $\limsup_{N\to\infty} a_N/b_N < \infty$.

We also define the following notation:

a. If X_n converges to X in distribution, we write $X_n \Rightarrow X$.

b. If X_n converges to X in probability, we write $X_n \to_p X$.

c. γ_d denotes the volume of the unit ball in $\mathbb{R}^d$.

d. For each $k \geq 1$, we define

$$\beta_k := \left(N\alpha^{(k-1)d} \prod_{i=1}^{k} \mu_i \right)^{-1/((k-1)d+k)}. \tag{1}$$

We will explain how β_k arises in Section 2.3.

e. σ_k denotes the first time a mutation of type k appears; a more rigorous definition of σ_k is given in Section 3.

The mutation rates $(\mu_i)_{i=1}^\infty$ and the rate of mutation spread α will depend on N, even though this dependence is not recorded in the notation. Throughout this paper we will assume that the mutation rates $(\mu_i)_{i=1}^\infty$ are asymptotically increasing, i.e.

$$\mu_1 \lesssim \mu_2 \lesssim \mu_3 \lesssim \mu_4 \lesssim \mu_5 \lesssim \cdots \tag{2}$$

In sections 2.1-2.4 we summarize results on the asymptotic distribution of σ_k, the first time a mutation of type k appears, assuming (2).

2.1 Theorem 1: low mutation rates

Assume

$$\mu_1 \ll \frac{\alpha}{N^{(d+1)/d}} \text{ and } \frac{\mu_i}{\mu_1} \to c_i \in (0, \infty] \text{ for all } i \in \{1, ..., k\}.$$

The first time a mutation of type 1 appears is exponentially distributed with rate $N\mu_1$. The maximal distance between any two points on the torus $\mathcal{T} = [0, L]^d$ is $\sqrt{d}L/2$. Also note that $L = N^{1/d}$ where N is the volume of $\mathcal{T}$. Subsequently, once the first type 1 mutation appears, it will spread to the entire torus in time $\sqrt{d}L/(2\alpha) = \sqrt{d}N^{1/d}/(2\alpha)$. Hence the time required for a type 1 mutation to fixate once it has first appeared is much shorter than σ_1 precisely when $N^{1/d}/\alpha \ll 1/(N\mu_1)$, which is equivalent to $\mu_1 \ll \alpha/N^{(d+1)/d}$.

Now because of the second assumption $\mu_i/\mu_1 \to c_i \in (0, \infty]$, mutations of types $i \in \{2, ..., k\}$ appear at least as fast as the first mutation. If $c_i < \infty$, then the waiting times σ_1 and $\sigma_i - \sigma_{i-1}$ are on the same order of magnitude. Because we have $\sigma_1 \sim \text{Exponential}(N\mu_1 c_1)$, it follows $\sigma_i - \sigma_{i-1}$ is also exponentially distributed and that $\sigma_i - \sigma_{i-1} \sim \text{Exponential}(N\mu_1 c_i)$. Otherwise, if $c_i = \infty$, then the first type i mutation appears so quickly that its waiting time $\sigma_i - \sigma_{i-1}$ is negligible as $N \to \infty$. Putting everything together gives us the following theorem.

Theorem 1. *Suppose $\mu_1 \ll \dfrac{\alpha}{N^{(d+1)/d}}$. Suppose that for all $i \in \{1, ..., k\}$, we have*

$$\frac{\mu_i}{\mu_1} \to c_i \in (0, \infty].$$

Let $W_1, ..., W_k$ be independent random variables with $W_i \sim \text{Exponential}(c_i)$ if $c_i < \infty$ and $W_i = 0$ if $c_i = \infty$. Then

$$N\mu_1\sigma_k \Rightarrow W_1 + \cdots + W_k.$$

We note that in the setting of asymptotically increasing mutation rates (i.e. assumption (2)), our result is slightly more general than Theorem 1 in [5]. Indeed, setting all mutation rates equal (i.e. $\mu_i = \mu$ for all i), we have $c_i = 1$ for all i, and we get $N\mu_1\sigma_k = N\mu\sigma_k \Rightarrow \text{Gamma}(k, 1)$, which is Case 1, Section 3 in [5]. However, Theorem 1 above is only a very slight variant of Theorem 1 of [5].

Figure 1 below illustrates that once a type i mutation appears, it quickly fills up the whole torus, and then a type $i + 1$ mutation occurs.
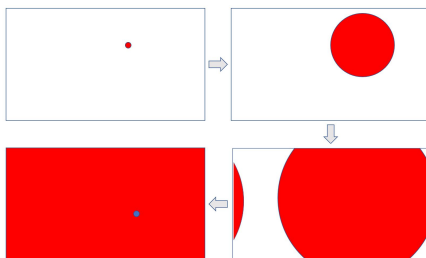

Figure 1. The transition from type i mutations (in red) to type $i + 1$ mutations (in blue).

2.2 Theorem 2: type $j \geq 2$ mutations occur rapidly after σ_1

Assume

$$\mu_1 \gg \frac{\alpha}{N^{(d+1)/d}} \text{ and } \mu_2 \gg \frac{(N\mu_1)^{d+1}}{\alpha^d}.$$

In contrast to Theorem 1, we have $\mu_1 \gg \alpha/N^{(d+1)/d}$. This means that the time it takes for type 1 mutations to fixate in the torus is longer than σ_1. As a result, there will be many balls of type 1 mutations before the entire torus is at least type 1.

We note that the $k = 2$ case was proven by Durrett, Foo, and Leder in Theorem 3 of [4]. We will be using a similar proof strategy, where we find a lower bound on how fast type $j \in \{2, ..., k\}$ mutations spread, using a single ball grown by a type $j - 1$ mutation. We first explain the $j = 2$ case. Focusing on a single ball grown by a type 1 mutation, the probability that the first type 2 mutation appears in that ball before time t is

$$1 - \exp\left(-\int_0^t \mu_2 \gamma_d(\alpha r)^d dr\right) = 1 - \exp\left(-\frac{\gamma_d}{d+1}\mu_2\alpha^d t^{d+1}\right). \tag{3}$$

It follows that the first time a type 2 mutation occurs in a single type 1 ball by time t is on the order of $(\mu_2\alpha^d)^{-1/(d+1)}$. Because there will be at least one ball of type 1, it follows that $\sigma_2 - \sigma_1$ is at least as fast as $(\mu_2\alpha^d)^{-1/(d+1)}$. Hence, whenever $(\mu_2\alpha^d)^{-1/(d+1)} \ll 1/(N\mu_1)$, which is equivalent to the second assumption, it follows that $\sigma_2 - \sigma_1$ is much quicker than σ_1. From this heuristic, we show that $N\mu_1(\sigma_2 - \sigma_1) \to_p 0$.

Note that in order to use equation (3) in the above argument, we need the volume of a single type 1 ball to equal the the volume of an Euclidean ball with the same radius, i.e. we require the type 1 ball to not start overlapping itself. This is true exactly when, for all times $r \in [0, t]$ in (3), the radius of a single type 1 ball at time r does not exceed $\sup\{|x - y| : x, y \in \mathcal{T}\} = \sqrt{d}N^{1/d}/2$. In other words, we need $\alpha r \leq \sqrt{d}N^{1/d}/2$. Rescaling time by a factor of $1/(N\mu_1)$, we want $\alpha r/(N\mu_1) \ll \sqrt{d}N^{1/d}/2$, which is satisfied due to the first assumption.

Now consider $j \in \{3, ..., k\}$. Repeating the same reasoning with types $j - 1, j$ in place of types $1, 2$, we see that $\sigma_j - \sigma_{j-1}$ is much quicker than σ_1 when $(\mu_j\alpha^d)^{-1/(d+1)} \ll 1/(N\mu_1)$, or equivalently $\mu_j \gg (N\mu_1)^{d+1}/\alpha^d$. However, this follows from our second assumption, because of (2). Hence we also have $N\mu_1(\sigma_j - \sigma_{j-1}) \to_p 0$. Putting everything together, when N is large,

$$N\mu_1\sigma_k = N\mu_1\sigma_1 + N\mu_1(\sigma_2 - \sigma_1) + \cdots + N\mu_1(\sigma_k - \sigma_{k-1}) \approx N\mu_1\sigma_1.$$

This gives us the following theorem.

Theorem 2. *Suppose that* $\mu_1 \gg \dfrac{\alpha}{N^{(d+1)/d}}$ *and* $\mu_2 \gg \dfrac{(N\mu_1)^{d+1}}{\alpha^d}$. *Then*

$$N\mu_1\sigma_k \Rightarrow W$$

where $W \sim Exponential(1)$.

A pictorial representation is given in Figure 2, where the decreasing nested circles correspond to mutations of types $1, ..., k$ for $k = 4$.
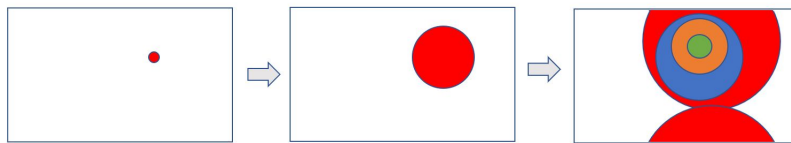
Figure 2: Once the first type 1 mutation (red) appears, the type 2 (blue), type 3 (orange), and type 4 (green) mutations all happen quickly.

2.3 Theorem 3: type $j \in \{1, ..., k-1\}$ mutations appear many times

Assume

$$\mu_1 \gg \frac{\alpha}{N^{(d+1)/d}} \text{ and } \mu_k \ll \frac{1}{\alpha^d \beta_{k-1}^{d+1}}.$$

As in Theorem 2, the first assumption ensures that σ_1 is shorter than the time it takes for type 1 mutations to fixate once they appear. The second assumption ensures that all mutations of types up to k do not appear too quickly, so that we are not in the setting of Theorem 2 (note that when $k = 2$, we have $\beta_{k-1} = (N\mu_1)^{-1}$, and the second assumption reduces to $\mu_2 \ll (N\mu_1)^{d+1}/\alpha^d$). In this case, there will be many small balls of type $j - 1$ before the jth mutation, which allows us to approximate the total volume of type $j - 1$ regions with its expectation.

The proof strategy in Theorem 3 will mostly follow Theorem 10 of [5], where we approximate the expectation of regions of at least type $j - 1$ with a deterministic function.

To explain what happens in this case, we repeat a derivation from [5]. We want to define an approximation $v_j(t)$ to the total volume of regions with at least j mutations at time t. We set $v_0(t) \equiv N$. Next, let $t > 0$. For times $r \in [0, t]$, type j mutations occur at rate $\mu_j v_{j-1}(r)$, and these type j mutations each grow into a ball of size $\gamma_d(\alpha(t - r))^d$ by time t. Then we write

$$v_j(t) = \int_0^t (\text{rate of type } j\text{th mutation at time } r)(\text{volume of ball from each type } j \text{ mutation})dr \tag{4}$$

$$= \int_0^t \mu_j v_{j-1}(r)\gamma_d(\alpha(t - r))^d dr.$$

Note that the informal equation (4) defining the approximation $v_j(t)$ is valid because we have many mostly non-overlapping balls of type $j - 1$. In [5] it is shown using induction that

$$v_j(t) = \frac{\gamma_d^j(d!)^j}{(j(d+1))!}\Big(\prod_{i=1}^j \mu_i\Big)N\alpha^{jd}t^{j(d+1)}$$

which gives us the approximation

$$\mathbb{P}(\sigma_k > t) \approx \exp\Big(-\int_0^t \mu_k v_{k-1}(r)dr\Big) = \exp\Big(-\frac{\gamma_d^{k-1}(d!)^{k-1}}{((k-1)d+k)!}\Big(\prod_{i=1}^k \mu_i\Big)N\alpha^{(k-1)d}t^{(k-1)d+k}\Big).$$

It follows that if we define β_k as in (1), then we have the following result.

Theorem 3. *Let $k \geq 2$. Suppose that $\mu_1 \gg \dfrac{\alpha}{N^{(d+1)/d}}$ and that $\mu_k \ll \dfrac{1}{\alpha^d \beta_{k-1}^{d+1}}$. Then*

$$\lim_{N\to\infty} \mathbb{P}(\sigma_k > \beta_k t) = \exp\Big(-\frac{\gamma_d^{k-1}(d!)^{k-1}}{((k-1)d+k)!}t^{(k-1)d+k}\Big).$$

When we have equal mutation rates (i.e. $\mu_i = \mu$ for all i), the conclusion above is the same as that of the third statement of Theorem 10 in [5]. Theorem 3 is illustrated in Figure 3, for $k = 3$.
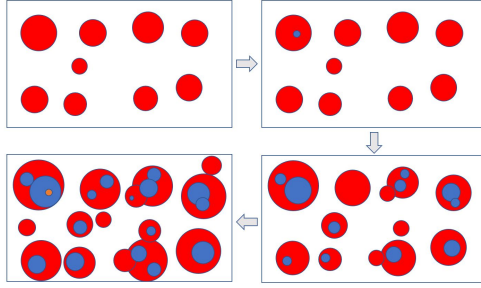
6

Figure 3: Mutations of types 1, 2, 3 colored in red, blue, and orange respectively.

2.4 Theorem 4: an intermediate case between Theorems 2 and 3

Assume

$$\mu_1 \gg \frac{\alpha}{N^{(d+1)/d}}.$$

We first define

$$l := \max\left\{ j \geq 2 : \mu_j \ll \frac{1}{\alpha^d \beta_{j-1}^{d+1}} \right\}. \tag{5}$$

Using (2) and Lemma 2 in Section 3.3, we in fact have

$$l = \max\left\{ j \geq 2 : \mu_2 \ll \frac{1}{\alpha^d \beta_1^{d+1}}, \mu_3 \ll \frac{1}{\alpha^d \beta_2^{d+1}}, ..., \mu_j \ll \frac{1}{\alpha^d \beta_{j-1}^{d+1}} \right\} \tag{6}$$

(Note that the definition of l in (5) omits the possibility $l = 1$, since β_0 is undefined. However, if in place of (5) we instead define $l := 1$, then Theorem 4 below is just Theorem 2.) If $l = -\infty$, then in particular $\mu_2 \ll 1/(\alpha^d \beta_1^{d+1})$, which is equivalent to $\mu_2 \ll (N\mu_1)^{d+1}/\alpha^d$. Then there are various possible behaviors for the type 2 mutations, inducing many edge cases which we do not pursue. On the other hand if $l \in \{k, k+1, ...\} \cup \{\infty\}$, then by (6) we have $\mu_k \ll 1/(\alpha^d \beta_{k-1}^{d+1})$, in which case Theorem 3 applies. Hence we assume $l \in \{2, ..., k-1\}$ and that

$$\mu_{l+1} \gg \frac{1}{\alpha^d \beta_l^{d+1}}. \tag{7}$$

The situation in Theorem 4 is a hybrid of Theorems 2 and 3. A mutation of type $j \in \{1, ..., l-1\}$ takes a longer time to fixate in the torus than interarrival time $\sigma_j - \sigma_{j-1}$. As a result, if $j \in \{2, ..., l\}$, there will be many mostly nonoverlapping balls of type $j-1$ before time σ_j. Using this fact, we proceed as in Theorem 3 and find $\lim_{N \to \infty} \mathbb{P}(\sigma_l > \beta_l t)$. Next, our assumption in (7) places us in the regime of Theorem 2; all mutations of types $l+1, ..., k$ happen so quickly that for all $\epsilon > 0$ we have $\mathbb{P}(\sigma_k - \sigma_l > \beta_l \epsilon) \to 0$. Then combining these two results yields the following result.

Theorem 4. *Suppose that* $\mu_1 \gg \dfrac{\alpha}{N^{(d+1)/d}}$, $l \in \{2, ..., k-1\}$, *and* $\mu_{l+1} \gg \dfrac{1}{\alpha^d \beta_l^{d+1}}$. *Then*

$$\mathbb{P}(\sigma_k > \beta_l t) \to \exp\left(-\frac{\gamma_d^{l-1}(d!)^{l-1}}{(d(l-1)+l)!} t^{d(l-1)+l} \right).$$

In pictures, Theorem 4 looks like Figure 3 for mutations up to type l. Then once the first type l mutation appears and spreads in a circle, all the subsequent mutations become nested within that circle, similar to Figure 2.

7

3 Proofs of Limit Theorems

In this section, we will prove the results stated in Section 2. We begin by introducing the structure of the torus $\mathcal{T} = [0, L]^d$, and will follow the notation of [5]. We define a pseudometric on the closed interval $[0, L]$ by

$$d_L(x, y) := \min\{|x - y|, L - |x - y|\}.$$

The d-dimensional torus of side length L will be denoted by $\mathcal{T} = [0, L]^d$. For $x = (x^1, ..., x^d) \in \mathcal{T}$ and $y = (y^1, ..., y^d) \in \mathcal{T}$ we define a pseudometric by

$$|x - y| := \sqrt{\sum_{i=1}^{d} d_L(x^i, y^i)^2}.$$

The torus should really be thought of as $\mathcal{T}$ modulo the equivalence relation $x \sim y$ iff $|x - y| = 0$, or more simply $\mathcal{T} = (\mathbb{R} \mod L)^d$. However, we will continue to write $\mathcal{T} = [0, L]^d$, keeping in mind that certain points are considered the same via the equivalence relation defined above. It will be useful to observe the following:

1. We have $d_L(x, y) \leq L/2$ for all $x, y \in [0, L]$. As a result,

$$\sup\{|x - y| : x, y \in \mathcal{T}\} \leq \sqrt{\sum_{i=1}^{d} \left(\frac{L}{2}\right)^2} = \frac{\sqrt{d}L}{2}. \tag{8}$$

2. Therefore, once a mutation of type j appears, the entire torus will become type j in time less than

$$\frac{\text{maximal distance between any } x, y \in \mathcal{T}}{\text{rate of mutation spread per unit time}} = \frac{\sqrt{d}L}{2\alpha}. \tag{9}$$

We use $|A|$ to denote the Lebesgue measure of some subset A (of $\mathcal{T}$ or $\mathcal{T} \times [0, \infty)$), so that $N = L^d = |\mathcal{T}|$ is the torus volume. Each $x \in \mathcal{T}$ at time t has a mutation of type $k \in \{0, 1, 2, ...\}$, which we denote by $T(x, t)$. The set of type i sites is defined by

$$\chi_i(t) := \{x \in \mathcal{T} : T(x, t) = i\}$$

The set of points whose type is at least i is defined by

$$\psi_i(t) := \{x \in \mathcal{T} : T(x, t) \geq i\} = \bigcup_{j=i}^{\infty} \chi_j(t).$$

At time t, we denote the total volume of type i sites by $X_i(t) := |\chi_i(t)|$, and the total volume of sites with type at least i by $Y_i(t) := |\psi_i(t)|$.

We now explicitly describe the construction of the process which gives rise to mutations in the torus. We will model mutations as random space-time points $(x, t) \in \mathcal{T} \times [0, \infty)$. Let $(\Pi_k)_{k=1}^{\infty}$ be a sequence of independent Poisson point processes on $\mathcal{T} \times [0, \infty)$, where Π_k has intensity μ_k. That is, for any space-time region $A \subseteq \mathcal{T} \times [0, \infty)$, the probability that A contains j points of type k is

$$e^{-\mu_k|A|} \frac{(\mu_k|A|)^j}{j!}.$$

8

Each $(x, t) \in \Pi_k$ is a space-time point at which $x \in \mathcal{T}$ can acquire a kth mutation at t. We say that x mutates to type k at time t precisely when $x \in \chi_{k-1}(t)$ and $(x, t) \in \Pi_k$. Once an individual obtains a type k mutation, it spreads the type k mutations outward in a ball at rate α. Here α depends only on N, the volume of the torus.

In the following proofs, we will assume that the mutation rates are asymptotically increasing, in the sense of assumption (2). With notation as defined above, the first time a type k mutation appears in the torus can be expressed as $\sigma_k = \inf\{t > 0 : Y_k(t) > 0\}$.

3.1 Proof of Theorem 1

We prove Theorem 1, which gives a result on the asymptotic distribution of σ_k assuming only that the first mutation fixates faster than σ_1. That is, once the first mutation appears, it spreads quickly to the entire torus. We note that the proof of Theorem 1 uses Theorem 1 of [5] by Foo, Leder, and Schweinsberg to calculate $N\mu_1\sigma_r$ where r is the maximal integer ($\leq k$) with $\mu_r \lesssim \mu_1$. We restate Theorem 1 in [5] below as Theorem A.

In the setting of (2), Theorem 1 is very similar to Theorem A when $j = 1$. However, Theorem 1 differs from Theorem A in that we only require the first mutation to have low rate (i.e. $\mu_1 \ll \alpha/N^{(d+1)/d}$), instead of requiring all $\mu_1, ..., \mu_{k-1}$ to be small relative to $\alpha/N^{(d+1)/d}$.

Theorem A. *Suppose $\mu_i \ll \alpha/N^{(d+1)/d}$ for $i \in \{1, ..., k-1\}$. Suppose there exists $j \in \{1, ..., k\}$ such that $\mu_j \ll \alpha/N^{(d+1)/d}$ and*

$$\frac{\mu_i}{\mu_j} \to c_i \in (0, \infty] \text{ for all } i \in \{1, ..., k\}.$$

Let $W_1, ..., W_k$ be independent random variables such that W_i has an exponential distribution with rate parameter c_i if $c_i < \infty$ and $W_i = 0$ if $c_i = \infty$. Then

$$N\mu_j\sigma_k \Rightarrow W_1 + \cdots + W_k.$$

Proof of Theorem 1. Let $r := \max\{j \in \{1, ..., k\} : \mu_j \lesssim \mu_1\}$. Then for any $j \in \{1, ..., r\}$,

$$0 \leq \lim_{N\to\infty} \frac{\mu_j}{\alpha/N^{(d+1)/d}} \leq \left(\lim_{N\to\infty} \frac{\mu_r}{\mu_1} \right) \left(\lim_{N\to\infty} \frac{\mu_1}{\alpha/N^{(d+1)/d}} \right) = c_r \cdot 0 = 0$$

and $\mu_j \ll \alpha/N^{(d+1)/d}$ for all $1 \leq j \leq r$. By Theorem A, we have

$$N\mu_1\sigma_r \Rightarrow W_1 + \cdots + W_r.$$

If $r = k$, then the conclusion follows. Otherwise, $r \leq k - 1$, and by maximality of r and equation (2), we have for all $l \in \{r+1, ..., k\}$

$$\frac{\mu_l}{\mu_1} \to \infty.$$

Then the result follows if we show

$$N\mu_1(\sigma_k - \sigma_r) \to_p 0.$$

We have

$$0 \leq N\mu_1(\sigma_k - \sigma_r) = N\mu_1 \sum_{j=r}^{k-1} (\sigma_{j+1} - \sigma_j). \tag{10}$$

9

We will find an upper bound for the right-hand side of (10). For $i \geq 1$ let $t_i = \inf\{t > 0 : Y_i(t) = N\}$ be the first time which every point in $\mathcal{T}$ is of at least type i. Define $\hat{t}_i := t_i - \sigma_i$, which is the time elapsed between σ_i and when mutations of type i fixate in the torus. Also define $\hat{\sigma}_i = \inf\{t > 0 : \Pi_i \cap (\mathcal{T} \times [t_{i-1}, t]) \neq \varnothing\}$, which is the first time there is a potential type i mutation after t_{i-1}. Observe that because we always have $\sigma_j \leq \hat{\sigma}_j$,

$$
\begin{aligned}
\sigma_{j+1} - \sigma_j &\leq \hat{\sigma}_{j+1} - \sigma_j \\
&= \hat{\sigma}_{j+1} - \sigma_j + t_j - t_j \\
&= \hat{t}_j + (\hat{\sigma}_{j+1} - t_j).
\end{aligned}
$$

Also observe that by (9), we have $\hat{t}_j \leq \sqrt{d} N^{1/d}/(2\alpha)$. Subsequently, the right-hand side of (10) has upper bound

$$
N\mu_1 \Big(\sum_{j=r}^{k-1} \hat{t}_j + \sum_{j=r}^{k-1} (\hat{\sigma}_{j+1} - t_j) \Big) \leq N\mu_1 (k - r) \frac{\sqrt{d} N^{1/d}}{2\alpha} + N\mu_1 \sum_{j=r}^{k-1} (\hat{\sigma}_{j+1} - t_j).
$$

The result follows if the right-hand side of the above converges to 0 in probability. The first term tends to zero because $\mu_1 \ll \alpha/N^{(d+1)/d}$. The second term tends to zero because $\hat{\sigma}_{j+1} - t_j \sim$ Exponential$(N\mu_{j+1})$, so $N\mu_1(\hat{\sigma}_{j+1} - t_j) \sim$ Exponential$(\mu_{j+1}/\mu_1) \to_p 0$ as $N \to \infty$. $\qquad \square$

3.2 Proof of Theorem 2

Lemma 1. *Let t_N be a random time that is $\sigma(\Pi_1, ..., \Pi_{j-1})$-measurable, depends on N, and satisfies $t_N \geq \sigma_{j-1}$. Then we have*

$$
\mathbb{P}(\sigma_j > t_N) = \mathbb{E}\left[\exp\left(-\int_{\sigma_{j-1}}^{t_N} \mu_j Y_{j-1}(s) ds \right) \right].
$$

Proof. Write $\mathcal{G} := \sigma(\Pi_1, ..., \Pi_{j-1})$. Define the $\mathcal{G}$-measurable set

$$
A := \{(x, r) \in \psi_{j-1}(r) \times [\sigma_{j-1}, t_N]\}.
$$

Conditioned on the σ-field $\mathcal{G}$, the event $\{\sigma_j > t_N\}$ occurs precisely when $\Pi_j \cap A = \varnothing$. It follows that if $X \sim$ Poisson$(\mu_j|A|)$, then

$$
\mathbb{P}(\sigma_j > t_N | \mathcal{G}) = \mathbb{P}(X = 0 | \mathcal{G}) = \exp\left(-\int_{\sigma_{j-1}}^{t_N} \mu_j Y_{j-1}(s) ds \right).
$$

Taking expectations of both sides finishes the proof. $\qquad \square$

Proof of Theorem 2. Write $N\mu_1\sigma_k$ as a telescoping sum

$$
N\mu_1\sigma_k = N\mu_1\sigma_1 + \sum_{j=2}^{k} N\mu_1(\sigma_j - \sigma_{j-1}).
$$

We have $N\mu_1\sigma_1 \sim$ Exponential(1). Hence it suffices to show that for each $j \geq 2$, the random variable $N\mu_1(\sigma_j - \sigma_{j-1})$ converges in probability to zero. Let $t > 0$. Then by Lemma 1,

$$
\begin{aligned}
\mathbb{P}(N\mu_1(\sigma_j - \sigma_{j-1}) > t) &= \mathbb{P}\Big(\sigma_j > \frac{t}{N\mu_1} + \sigma_{j-1} \Big) \\
&= \mathbb{E}\left[\exp\left(-\int_{\sigma_{j-1}}^{t/(N\mu_1)+\sigma_{j-1}} \mu_j Y_{j-1}(s) ds \right) \right].
\end{aligned}
$$

We want to show that the term on the right-hand side tends to zero. By the dominated convergence theorem, it suffices to show that as $N \to \infty$

$$\int_{\sigma_{j-1}}^{t/(N\mu_1)+\sigma_{j-1}} \mu_j Y_{j-1}(s)ds \to \infty \text{ a.s.}$$

Notice that because $\mu_1 \gg \alpha/N^{(d+1)/d}$, for all sufficiently large N we have $t/(N\mu_1) \le \sqrt{d}N^{1/d}/(2\alpha)$. By (9), we deduce that $t/(N\mu_1)$ does not exceed the time it takes for type $j-1$ mutations to fixate. Hence we have $Y_{j-1}(s) \ge \gamma_d\alpha^d(s-\sigma_{j-1})^d$ for $s \in [\sigma_{j-1}, \sigma_{j-1}+t/(N\mu_1)]$, and that

$$\int_{\sigma_{j-1}}^{t/(N\mu_1)+\sigma_{j-1}} \mu_j Y_{j-1}(s)ds \ge \int_{\sigma_{j-1}}^{t/(N\mu_1)+\sigma_{j-1}} \mu_j\gamma_d\alpha^d(s-\sigma_{j-1})^d ds$$

$$= \int_0^{t/(N\mu_1)} \mu_j\gamma_d\alpha^d u^d du$$

$$= \frac{\mu_j\gamma_d\alpha^d}{d+1}\left(\frac{t}{N\mu_1}\right)^{d+1}$$

It remains to show

$$\frac{\mu_j\gamma_d\alpha^d}{d+1}\left(\frac{t}{N\mu_1}\right)^{d+1} \to \infty \text{ as } N \to \infty.$$

For the above to hold, it suffices to have $\mu_j \gg (N\mu_1)^{d+1}/\alpha^d$, which holds due to the second assumption in the theorem and equation (2). This finishes the proof. $\square$

3.3 Proof of Theorem 3

We recall the definition of β_k as in (1) of Section 2. In the setting of Theorem 3, β_k is the order of magnitude of the time it takes for the kth mutation to appear.

Much of the proof of Theorem 3 will rely on Lemma 9 of [5], which approximates a monotone stochastic process by a deterministic function under a certain time-scaling. In order to apply this lemma, it is important to ensure that $Y_k(t)$ is well-approximated by its expectation, which is Lemma 8 of the same paper.

Before proving Theorem 3, we state several lemmas, some of which are from [5]. First, we need to ensure that the last assumption $\mu_k\alpha^d\beta_{k-1}^{d+1} \to 0$ in Theorem 3 implies $\mu_k\alpha^d\beta_k^{d+1} \to 0$, so that we are able to use part 2 of Lemma 5 to approximate $Y_{k-1}(\beta_k t)$ by its expectation.

Lemma 2. *For $k \ge 2$, $\mu_k \ll \dfrac{1}{\alpha^d\beta_k^{d+1}}$ if and only if $\mu_k \ll \dfrac{1}{\alpha^d\beta_{k-1}^{d+1}}$.*

Proof. By direct calculation

$$\mu_k \ll \frac{1}{\alpha^d\beta_k^{d+1}} \iff \mu_k \ll \frac{1}{\alpha^d}\left(N\alpha^{(k-1)d}\prod_{i=1}^k \mu_i\right)^{\frac{d+1}{(k-1)d+k}}$$

$$\iff \mu_k^{(k-1)d+k} \ll \frac{1}{\alpha^{d[(k-1)d+k]}}\left(N\alpha^{(k-1)d}\prod_{i=1}^k \mu_i\right)^{d+1}$$

$$\iff \mu_k^{(k-1)d+k} \ll \frac{\alpha^{d(d+1)(k-1)}}{\alpha^{d[(k-1)d+k]}}N^{d+1}\left(\prod_{i=1}^k \mu_i\right)^{d+1}$$

11

$$\Longleftrightarrow \mu_k^{(k-1)d+k} \ll \frac{1}{\alpha^d} N^{d+1} \Big(\prod_{i=1}^{k} \mu_i \Big)^{d+1}$$

$$\Longleftrightarrow \mu_k^{(k-2)d+(k-1)} \ll \frac{1}{\alpha^d} N^{d+1} \Big(\prod_{i=1}^{k-1} \mu_i \Big)^{d+1}$$

$$\Longleftrightarrow \mu_k^{(k-2)d+(k-1)} \ll \frac{\alpha^{d(d+1)(k-2)}}{\alpha^{d[(k-2)d+(k-1)]}} N^{d+1} \Big(\prod_{i=1}^{k-1} \mu_i \Big)^{d+1}$$

$$\Longleftrightarrow \mu_k^{(k-2)d+(k-1)} \ll \frac{1}{\alpha^{d[(k-2)d+(k-1)]}} \Big(N \alpha^{(k-2)d} \prod_{i=1}^{k-1} \mu_i \Big)^{d+1}$$

$$\Longleftrightarrow \mu_k \ll \frac{1}{\alpha^d \beta_{k-1}^{d+1}}.$$

$\square$

Second, we also need Lemma 9 from [5], which gives necessary conditions under which a monotone stochastic process is well-approximated by a deterministic function.

Lemma 3. *Suppose, for all positive integers N, $(Y_N(t), t \geq 0)$ is a nondecreasing stochastic process with finite mean for all $t > 0$. Assume there exist sequences of positive numbers $(\nu_N)_{N=1}^\infty$ and $(s_N)_{N=1}^\infty$ and a continuous nondecreasing function $g > 0$ such that for all $t > 0$ and $\epsilon > 0$, we have*

$$\lim_{N \to \infty} \mathbb{P}(|Y_N(s_N t) - \mathbb{E}[Y_N(s_N t)]| > \epsilon \mathbb{E}[Y_N(s_N t)]) = 0 \tag{11}$$

and

$$\lim_{N \to \infty} \frac{1}{\nu_N} \mathbb{E}[Y_N(s_N t)] = g(t). \tag{12}$$

Then for all $\epsilon > 0$ and $\delta > 0$, we have

$$\lim_{N \to \infty} \mathbb{P}(\nu_N g(t)(1 - \epsilon) \leq Y_N(s_N t) \leq \nu_N g(t)(1 + \epsilon) \ \text{for all} \ t \in [\delta, \delta^{-1}]) = 1.$$

Third, we will state a criterion which guarantees that for fixed $t > 0$, the probability $\mathbb{P}(\sigma_k > \beta_k t)$ converges to a deterministic function as $N \to \infty$.

Lemma 4. *For a continuous nonnegative function g, a positive sequence $(\nu_N)_{N=1}^\infty \subseteq \mathbb{R}^+$, and $\delta, \epsilon > 0$, define the event*

$$B_N^{k-1}(\delta, \epsilon, g, \nu_N) = \{g(u)(1 - \epsilon)\nu_N \leq Y_{k-1}(\beta_k u) \leq g(u)(1 + \epsilon)\nu_N, \ \text{for all} \ u \in [\delta, \delta^{-1}]\}.$$

If $(\nu_N)_{N=1}^\infty$ and g are chosen such that $\lim_{N \to \infty} \nu_N \beta_k \mu_k$ exists and $\lim_{N \to \infty} \mathbb{P}(B_N^{k-1}(\delta, \epsilon, g, \nu_N)) = 1$, then we have

$$\lim_{N \to \infty} \mathbb{P}(\sigma_k > \beta_k t) = \lim_{N \to \infty} \exp\Big(- \nu_N \beta_k \mu_k \int_0^t g(u) du \Big).$$

Proof. Reasoning as in the proof of Theorem 10 from [5], we have the upper and lower bounds

$$\mathbb{P}(\sigma_k > \beta_k t) \leq \exp\Big(- \mu_k \beta_k \nu_N (1 - \epsilon) \int_\delta^t g(u) du \Big) + \mathbb{P}(B_N^{k-1}(\delta, \epsilon, g, \nu_N)^c)$$

$$\mathbb{P}(\sigma_k > \beta_k t) \geq \mathbb{P}(B_N^{k-1}(\delta, \epsilon, g, \nu_N)) \exp\left(-\nu_N(1+\epsilon)\beta_k\mu_k \int_\delta^t g(u)du \right) - \frac{\gamma_d^{k-1}(d!)^{k-1}}{(d(k-1)+k)!} \delta^{d(k-1)+k}.$$

The above inequalities are equations (26) and (27) in [5], respectively. Taking $N \to \infty$ then $\epsilon, \delta \to 0$, we get the desired result. $\qquad\square$

Lastly, we need to approximate the expected volume of type k or higher regions, $\mathbb{E}[Y_k(t)]$, with a deterministic function, as well as making sure that $Y_k(t)$ is well-approximated by its expectation.

We remark that in Lemma 5 below, time t may depend on N. This is because in the proof of Theorem 3, we will rescale time by a factor of β_k, i.e. replace t with $\beta_k t$.

Lemma 5. *Fix a positive integer k. Suppose $\mu_j \alpha^d t^{d+1} \to 0$ for all $j \in \{1, ..., k\}$. Also suppose $t \ll N^{1/d}/(2\alpha)$. Then*

1. *Setting $v_k(t) := \dfrac{\gamma_d^k(d!)^k}{(k(d+1))!}\left(\prod_{i=1}^k \mu_i\right) N\alpha^{kd}t^{k(d+1)}$, we have $\mathbb{E}[Y_k(t)] \sim v_k(t)$.*

2. *If in addition we assume $\left(\prod_{i=1}^k \mu_i\right) N\alpha^{(k-1)d}t^{(k-1)d+k} \to \infty$, then for all $\epsilon > 0$,*

$$\lim_{N\to\infty} \mathbb{P}((1-\epsilon)\mathbb{E}[Y_k(t)] \leq Y_k(t) \leq (1+\epsilon)\mathbb{E}[Y_k(t)]) = 1.$$

Proof. The first and second statements are Lemmas 5 and 8 in [5], respectively. $\qquad\square$

Remark. Lemma 5 in [5] omits the necessary hypothesis $t \ll N^{1/d}/(2\alpha)$. This is needed in order for

$$\mathbb{E}[\Lambda_{k-1}(t)] = \mu_k\gamma_d\alpha^d \int_0^t \mathbb{P}(0 \in \psi_{k-1}(s))(t-s)^d ds$$

in equation (15) of [5] to hold, so that the inequality in the proof of Lemma 5 of [5]

$$\mathbb{P}(0 \in \psi_j(s)) \geq (1-\epsilon)\mathbb{E}[\Lambda_{j-1}(s)] \geq (1-\epsilon)\mu_j\gamma_d\alpha^d \int_0^s \mathbb{P}(0 \in \psi_{j-1}(u))(s-u)^d du$$

is valid. Note that the hypothesis $t \ll N^{1/d}/(2\alpha)$ is also necessary for Lemma 8 in [5], because its proof uses Lemma 5 in [5].

Lemma 6. *If $\mu_1 \gg \dfrac{\alpha}{N^{(d+1)/d}}$, then $\beta_l \ll \dfrac{N^{1/d}}{\alpha}$ for any $l \in \mathbb{N}$.*

Proof. By (2), we have $\mu_1, ..., \mu_l \gg \alpha/N^{(d+1)/d}$. Thus

$$\prod_{i=1}^l \mu_i \gg \frac{\alpha^l}{N^{l(1+1/d)}}.$$

On the other hand by simplifying,

$$\beta_l \ll \frac{N^{1/d}}{\alpha} \iff N\alpha^{(l-1)d}\prod_{i=1}^l \mu_i \gg \left(\frac{\alpha}{N^{1/d}}\right)^{(l-1)d+l}$$

$$\iff \prod_{i=1}^l \mu_i \gg \frac{\alpha^l}{N^{l(1+1/d)}}.$$

This proves the lemma. $\qquad\square$

Proof of Theorem 3. In view of Lemma 4, we will choose $(\nu_N)_{N=1}^{\infty}$ and a continuous nonnegative function g such that $\lim_{N\to\infty} \nu_N \beta_k \mu_k$ exists and $\mathbb{P}(B_N^{k-1}(\delta, \epsilon, g, \nu_N)) \to 1$ as $N \to \infty$. We set $\nu_N = 1/(\beta_k \mu_k)$, and as in the proof of Theorem 10 in [5], set

$$g = g_k(t) := \frac{\gamma_d^{k-1}(d!)^{k-1} t^{(k-1)(d+1)}}{((k-1)(d+1))!}.$$

A lengthy calculation shows that $\beta_k \mu_k v_{k-1}(\beta_k t) = g_k(t)$. On the other hand, by the last assumption in the theorem, we have $\mu_k \alpha^d \beta_{k-1}^{d+1} \to 0$. By Lemma 2, this is equivalent to $\mu_k \alpha^d \beta_k^{d+1} \to 0$. Because of (2), this implies that as $N \to \infty$

$$\mu_j \alpha^d (\beta_k t)^{d+1} \to 0$$

for all $j \in \{1, ..., k\}$. Also, because of Lemma 6, we have $\beta_k \ll N^{1/d}/(2\alpha)$. Hence the hypotheses of Lemma 5 are satisfied, and by the first result in Lemma 5 applied to $k-1$, it follows that $v_{k-1}(\beta_k t) \sim \mathbb{E}[Y_{k-1}(\beta_k t)]$, which implies

$$\lim_{N\to\infty} \beta_k \mu_k \mathbb{E}[Y_{k-1}(\beta_k t)] = \lim_{N\to\infty} \beta_k \mu_k v_{k-1}(\beta_k t) = g_k(t).$$

Hence, (12) of Lemma 3 is satisfied, and it remains to check (11). To this end, by the second result of Lemma 5, it suffices to show that

$$\Big(\prod_{i=1}^{k-1} \mu_i\Big) N\alpha^{(k-2)d} \beta_k^{(k-2)d+k-1} \to \infty.$$

A direct calculation gives

$$\Big(\prod_{i=1}^{k-1} \mu_i\Big) N\alpha^{(k-2)d} \beta_k^{(k-2)d+k-1} = \frac{1}{\mu_k \alpha^d \beta_k^{d+1}} \to \infty.$$

Therefore, Lemma 3 guarantees that $\mathbb{P}(B_N^{k-1}(\delta, \epsilon, g, \nu_N)) \to 1$ as $N \to \infty$. Then, Lemma 4 gives us

$$
\begin{aligned}
\lim_{N\to\infty} \mathbb{P}(\sigma_k > \beta_k t) &= \lim_{N\to\infty} \exp\Big(-\nu_N \beta_k \mu_k \int_0^t g_k(u)du\Big) \\
&= \exp\Big(-\int_0^t \frac{\gamma_d^{k-1}(d!)^{k-1} u^{(k-1)(d+1)}}{((k-1)(d+1))!} du\Big) \\
&= \exp\Big(-\frac{\gamma_d^{k-1}(d)!^{k-1}}{(d(k-1)+k)!} t^{d(k-1)+k}\Big)
\end{aligned}
$$

finishing the proof. $\square$

3.4 Proof of Theorem 4.

Now we turn to proving Theorem 4, which is a hybrid of Theorems 2 and 3. In particular, we assume that there is some $l \in \mathbb{N}$ such that the mutation rates $\mu_1, \mu_2, ..., \mu_l$ fall under the regime of Theorem 3, and all subsequent mutation rates $\mu_{l+1}, ..., \mu_k$ are large enough so that all mutations after the lth one occur quickly, as in Theorem 2.

Proof of Theorem 4. For ease of notation, set, for $j \in \mathbb{N}$ and $t \geq 0$,

$$f_j(t) := \exp\left(- \frac{\gamma_d^{j-1}(d)!^{j-1} t^{d(j-1)+j}}{(d(j-1)+j)!} \right)$$

For $\epsilon > 0$, we have the inequalities

$$\mathbb{P}(\sigma_l > \beta_l t) \leq \mathbb{P}(\sigma_k > \beta_l t) \leq \mathbb{P}(\sigma_l > \beta_l(t-\epsilon)) + \mathbb{P}(\sigma_k - \sigma_l > \beta_l \epsilon).$$

Taking $N \to \infty$ and using Theorem 3 (noting that $l \geq 2$), we have

$$f_l(t) \leq \lim_{N\to\infty} \mathbb{P}(\sigma_k > \beta_l t) \leq f_l(t-\epsilon) + \lim_{N\to\infty} \mathbb{P}(\sigma_k - \sigma_l > \beta_l \epsilon).$$

Since f_l is continuous, the result follows (by taking $\epsilon \to 0$) once we show that for each fixed $\epsilon > 0$

$$\lim_{N\to\infty} \mathbb{P}(\sigma_k - \sigma_l > \beta_l \epsilon) = 0.$$

Notice that we have

$$\{\sigma_k - \sigma_l > \beta_l \epsilon\} \subseteq \bigcup_{j=l}^{k-1} \left\{ \sigma_{j+1} - \sigma_j > \frac{\beta_l \epsilon}{k-l} \right\}$$

from which we deduce

$$\mathbb{P}(\sigma_k - \sigma_l > \beta_l \epsilon) \leq \sum_{j=l}^{k-1} \mathbb{P}\left(\sigma_{j+1} - \sigma_j > \frac{\beta_l \epsilon}{k-l} \right).$$

Replacing $\epsilon/(k-l) > 0$ with $\epsilon > 0$ without loss of generality, to finish the proof it is enough to show that for all $j \in \{l, ..., k-1\}$ we have

$$\mathbb{P}(\sigma_{j+1} - \sigma_j > \beta_l \epsilon) \to 0.$$

By Lemma 1, we have

$$\mathbb{P}(\sigma_{j+1} - \sigma_j > \beta_l \epsilon) = \mathbb{E}\left[\exp\left(- \int_{\sigma_j}^{\beta_l \epsilon + \sigma_j} \mu_{j+1} Y_j(s) ds \right) \right].$$

Hence, by the dominated convergence theorem, to show that $\mathbb{P}(\sigma_{j+1} - \sigma_j > \beta_l \epsilon) \to 0$, it suffices to show that

$$\int_{\sigma_j}^{\beta_l \epsilon + \sigma_j} \mu_{j+1} Y_j(s) ds \to \infty \text{ a.s. as } N \to \infty.$$

By Lemma 6 we have $\beta_l \ll N^{1/d}/\alpha$, so $\beta_l \epsilon \leq \sqrt{d} N^{1/d}/(2\alpha)$ for all large enough N. That is, $\beta_l \epsilon$ does not exceed the time it takes for the jth mutation to fixate. Hence we have the lower bound $Y_j(s) \geq \gamma_d \alpha^d (s - \sigma_j)^d$ for $s \in [\sigma_j, \sigma_j + \beta_l \epsilon]$, and

$$\int_{\sigma_j}^{\beta_l \epsilon + \sigma_j} \mu_{j+1} Y_j(s) ds \geq \int_{\sigma_j}^{\beta_l \epsilon + \sigma_j} \mu_{j+1} \gamma_d \alpha^d (s - \sigma_j)^d = \frac{\mu_{j+1} \gamma_d \alpha^d}{d+1} (\beta_l \epsilon)^{d+1}.$$

15

It remains to show that the right-hand side of the above term tends to infinity. By the second assumption in the theorem, we have $\mu_{l+1} \gg 1/(\alpha^d \beta_l^{d+1})$. Because of (2), we have $\mu_{j+1} \gg \mu_{l+1}$. Hence it follows that as $N \to \infty$

$$\frac{\mu_{j+1} \gamma_d \alpha^d}{d+1}(\beta_l \epsilon)^{d+1} \to \infty.$$

The proof is complete. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

References

[1] P. Armitage and R. Doll (1954). The age distribution of cancer and a multi-stage theory of carcinogenesis. *Brit. J. Cancer* **8**, 1-12.

[2] M. Bramson and D. Griffeath (1980). On the Williams-Bjerknes tumour growth model: II. *Math. Proc. Camb. Phil. Soc.* **88**, 339-357.

[3] M. Bramson and D. Griffeath (1981). On the Williams-Bjerknes tumour growth model: I. *Ann. Probab.* **9**, 173-185.

[4] R. Durrett, J. Foo, and K. Leder (2016). Spatial Moran models II: cancer initiation in spatially structured tissue. *J. Math. Biol.* **72**, 1369-1400.

[5] J. Foo, K. Leder, and J. Schweinsberg (2020). Mutation timing in a spatial model of evolution. *Stochastic Process. Appl.* **130**, 6388-6413.

[6] K. R. Loeb and L. A. Loeb (2000). Significance of multiple mutations in cancer. *Carcinogenesis* **21**, 379-385.

[7] M. J. Prindle, E. J. Fox, L. A. Loeb (2010). The mutator phenotype in cancer: molecular mechanisms and targeting strategies. *Curr Drug Targets* **11(10)**, 1296-1303.

[8] T. Williams and R. Bjerknes (1972). Stochastic model for abnormal clone spread through epithelial basal layer. *Nature* **236**, 19-21.