



PHD

## Accessibility Percolation and Evolutionary Dynamics in Varying Environments

Bartos, Thomas

*Award date:*  
2023

*Awarding institution:*  
University of Bath

[Link to publication](#)

### Alternative formats

If you require this document in an alternative format, please contact:  
[openaccess@bath.ac.uk](mailto:openaccess@bath.ac.uk)

Copyright of this thesis rests with the author. Access is subject to the above licence, if given. If no licence is specified above, original content in this thesis is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC-ND 4.0) Licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>). Any third-party copyright material present remains the property of its respective owner(s) and is licensed under its existing terms.

#### Take down policy

If you consider content within Bath's Research Portal to be in breach of UK law, please contact: [openaccess@bath.ac.uk](mailto:openaccess@bath.ac.uk) with the details. Your claim will be investigated and, where appropriate, the item will be removed from public view as soon as possible.

**Accessibility Percolation and  
Evolutionary Dynamics in Varying  
Environments**

submitted by

Thomas Bartos

for the degree of Doctor of Philosophy

of the

University of Bath

Department of Mathematical Sciences

December 2022



## **COPYRIGHT**

Attention is drawn to the fact that copyright of this thesis rests with the author. A copy of this thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that they must not copy it or use material from it except as permitted by law or with the consent of the author.

Access to this thesis in print or electronically is restricted until ..... (date)

Signed on behalf of the Doctoral College by .....

**Declaration of any previous submission of work**

The material presented here for examination for the award of a higher degree by research has not been incorporated into a submission for another degree.

Candidate's typed signature.....Thomas Bartos.....

**Declaration of authorship**

I am the author of this thesis, and the work described therein was carried out by myself personally, with the exception of Chapter 3 where 60% of the work (formulation of ideas) was carried out by other researchers.

Candidate's typed signature.....Thomas Bartos.....

## Summary

In this thesis I investigate various mathematical problems that are either connected to or loosely inspired by evolutionary biology. The main focus is on variants of existing models with a special emphasis on dynamic models or those in which the random quantities that define the model change over time. The first two chapters are concerned with accessibility percolation on tree-like graphs. I first look at a dynamic model on finite (but arbitrarily large) trees and then consider a static model on infinite (but locally finite) trees. For the dynamic model I consider two different scalings of the height of the tree, with the second scaling corresponding to a ‘critical window’ for the first scaling. For each of these scalings, I show that there exist times within an interval of polynomial size in the height of the tree, at which an accessible path exists (with probability converging to 1 as the height of the tree goes to infinity). For the infinite tree, a criterion is defined that quantifies the amount of branching in the tree, and this criterion is shown to determine a phase transition for the existence of an accessible path. In the second part of the thesis, I investigate variants of a model of adaptive dynamics, in which two mutants compete for fixation. The first variant involves an established type and an invading type, with the growth rate of the invading type subject to periodic switches between two values. I show that the length of time taken to invade is averaged provided the period length grows sufficiently slowly with respect to the carrying capacity  $K$ . If the period length decreases to zero with increasing  $K$ , the probability of invasion is itself averaged. The second variant involves two non-established types with one type initiated after a delay. The probability that either type reaches a size proportional to  $K$  is calculated for different initial starting sizes and it is shown that for fixed starting sizes, there is a phase transition in the duration of the allowed delay on the  $\log K$  timescale.

# Contents

<b>1</b>	<b>Introduction</b>	<b>6</b>
1.1	Accessibility Percolation . . . . .	6
1.1.1	Biological Motivation . . . . .	6
1.1.2	Accessibility Percolation on Regular Trees . . . . .	9
1.1.3	Accessibility Percolation on Infinite Trees . . . . .	15
1.2	Adaptive Dynamics . . . . .	18
1.2.1	General Theory and Motivation . . . . .	18
1.2.2	Invasion in a Periodic Environment . . . . .	21
1.2.3	Invasion with Delay . . . . .	24
<b>2</b>	<b>Dynamic Accessibility Percolation on Regular Trees</b>	<b>28</b>
2.1	Introduction . . . . .	28
2.2	Overview of Proofs . . . . .	29
2.3	Subcritical Cases . . . . .	30
2.4	Supercritical Case: Increasing paths exist at all times in $[0, 1]$ when $\alpha > \frac{1}{e}$	35
2.5	Critical case: Second moment bounds . . . . .	39
2.5.1	Cases $c(u, v) = 0$ and $c(u, v) = h$ . . . . .	42
2.5.2	Correlations over short time intervals: $ t - s  < \delta_n$ . . . . .	45
2.5.3	Correlations over long time intervals: $ t - s  \geq \delta_n$ . . . . .	47
2.6	Extending from the first $K$ levels . . . . .	61
<b>3</b>	<b>Accessibility Percolation on an Infinite Tree</b>	<b>76</b>
3.1	Introduction . . . . .	76
3.2	Electrical Networks . . . . .	77
3.3	Application to Accessibility Percolation on General Trees . . . . .	80
3.4	Spherically Symmetric Trees in the Critical Case . . . . .	84
3.5	Possible Extension to Dynamical Version . . . . .	87
<b>4</b>	<b>Invasion in a Periodic Environment</b>	<b>90</b>

4.1	Model . . . . .	90
4.2	Time to Reach a Level . . . . .	95
4.3	First Phase . . . . .	104
4.4	Second Phase . . . . .	115
4.5	Possible Extension to Random Environment . . . . .	123
4.6	Auxiliary Results . . . . .	126
<b>5</b>	<b>Invasion with Delay</b>	<b>131</b>
5.1	Probability of Invasion . . . . .	132
5.2	Possible Extension to Allometric Scaling . . . . .	144
<b>A</b>	<b>Appendix</b>	<b>146</b>
A.1	Stochastic Integral With Respect To Poisson Processes . . . . .	146
A.2	Itô's Formula for Jump Processes . . . . .	148
A.3	Factorial Branching Number on Spherically Symmetric Trees . . . . .	149



# Chapter 1

## Introduction

### 1.1 Accessibility Percolation

#### 1.1.1 Biological Motivation

The concept of a fitness landscape was originally introduced to model the genetic evolution of a biological population. The modern synthesis of evolutionary biology [22], which was an early framework combining the ideas of Darwinian natural selection with Mendel's discoveries of genetic inheritance, formulates the evolution of asexual populations as follows: all individuals in a population begin with the same set of genes (referred to as its genome), and over time each individual accumulates random changes in their set of genes, with the new genes referred to as mutations. On reproduction, a copy of the parent's genetic sequence is transmitted to their offspring, who in turn experience further mutations. The exact sequence of genes possessed by an individual is referred to as its genotype, and assuming an individual's genotype directly specifies that individual's traits<sup>1</sup>, one can assume that given a fixed environment, the genotype determines the individual's ability to survive and consequently the number of offspring produced. The environment thus applies a selective pressure to each individual, and genes which enhance an individual's ability to survive will be transmitted in greater numbers to the next generation. By this process, certain selectively advantageous genes will eventually be shared by all extant individuals of the population. Any such gene is said to have reached *fixation* in the population.

Note that in general a biological population is not composed of a single type, however

---

<sup>1</sup>An assumption not true in general as external environmental factors often have a significant effect.

under certain conditions a population could be considered genetically homogeneous at a particular locus (genomic segment). Such homogeneity is attained (or approximately attained) in the *strong selection–weak mutation regime*. This means the mutation rate is low compared to the timescale of fixation of beneficial mutations - a new mutation arising in an individual that confers higher fitness will fix in the population before a new mutation arises anywhere in the population. In this regime it is therefore reasonable to model the population as a single particle whose genotype takes values in some metric space.

The first approach in this direction is known as the *geometric model* and was put forward by Fisher [18]. In the geometric model, a population is represented as a  $k$ -vector  $\mathbf{x} \in \mathbb{R}^k$  for some  $k \in \mathbb{N}$  and the selective advantage or *fitness*  $f(\mathbf{x})$  of the population is inversely proportional to the distance from the origin; for example one could take as the fitness function  $f(\mathbf{x}) = 1/\|\mathbf{x}\|_2 = (\sum_{i=1}^k x_i^2)^{-\frac{1}{2}}$ . Each coordinate represents some continuous trait possessed by the organisms in the population. Fisher supposed that a population begins as a point some distance from the origin, and by random addition of vectors of varying length, a population changes its overall fitness. For such random additions  $\mathbf{y} \in \mathbb{R}^k$ , the point  $\mathbf{x} + \mathbf{y}$  represents the effect of a mutation which spreads to the entire population. The forces of natural selection mean that the closer  $\mathbf{x} + \mathbf{y}$  is to the origin, the more likely it is that the population will jump to this point, and thus the observed trajectory of a surviving population will be a path that over time moves towards the origin. This idea that the population under consideration evolves as a point in some metric space (or *landscape*) on which a function quantifying the fitness of an organism is defined, can be taken further. In particular, the landscape can be thought of as a finite graph with each of the vertices corresponding to genetic types and each of the edges representing a possible mutational step between genotypes. Each genotype has an associated fitness, which if viewed as a random quantity can be modelled by attaching a real-valued random variable to each vertex. The prototypical example of a graph representing a genetic type-space is the  $n$ -dimensional hypercube  $\mathbb{H}_n$ . The reason for this choice is that each of the vertices of the hypercube can be uniquely coded in a natural way for any genomic segment consisting of diallelic genes (genes coming in two types). Specifically, the dimension  $n$  of the hypercube represents the length of the genomic segment and each of the  $n$  sites in the segment can be occupied by one of two alleles, coded by 0 and 1. The vertex set is then in one-to-one correspondence with  $\{0, 1\}^n$ , the set of all binary strings of length  $n$ , and each edge links genotypes differing by a single mutation (the binary strings associated to the terminal vertices differ at a single position). We designate the root vertex of the graph to be the vertex coded by the string of all zeros (this could be chosen anywhere as the hypercube is

transitive, once this is fixed the coding of the remaining vertices is determined). The idea is that the root represents the ‘wild type’ and as the population moves away from root it accumulates new mutations, represented by a 0 changing to a 1 at a single site in the genotype as it traverses each edge. This more specific fitness landscape concept was originally introduced by Wright [38].

The hypercube geometry can however be replaced by a simpler tree-based geometry. As the dimension of the hypercube grows, each ‘half’ of the hypercube becomes better approximated by a tree. In particular, when the graph is directed away from a designated corner vertex, for all vertices at graph distance  $d$  from this vertex, the ratio of the in-degree to the out-degree is  $d/(n-d)$  which is less than 1 for  $1 \leq d < n/2$  and converges to 0 for  $d = o(n^\gamma)$  where  $\gamma < 1$ . Biologically, the approximation is also reasonable as the probability of mutational reversions is assumed to be low. This is one of the main motivations for studying a tree-like graph as our underlying fitness landscape in this thesis (along with the fact that it allows us to analyse a model with a phase transition whose behaviour is known at criticality - this will be explained below). The strong selection - weak mutation regime constrains the population as a single particle to move through the landscape along paths of increasing fitness. A path from the root to the terminal vertex of a tree with monotonically increasing fitness can thus be considered ‘accessible via evolution’, a concept first identified in a biological context by Weinreich, Watson and Chao [14]. Assuming that the fitness of each vertex is a random variable leads to the mathematical question of the probability that an accessible path exists in the underlying graph. This notion is referred to as *accessibility percolation* and was first studied by Nowak and Krug in [30]. We will simplify the situation by requiring the random variables giving the fitness values of the different genotypes to be independent and identically distributed<sup>2</sup>. However we will alter the setup by specifying that the fitness values are independently resampled at unit rate, leading to a dynamic fitness landscape. The interesting questions for this landscape then relate to the time intervals for which we will (or will not) see an accessible path with high probability. It should be noted that these problems are of primarily mathematical interest as a biologically realistic dynamic fitness landscape would have to incorporate correlations both in fitness and in the resampling times, which would make the analysis very complicated.

Nevertheless, there have been a number of works analysing the more biologically relevant aspects of dynamic fitness landscapes, including correlated landscapes. Many

---

<sup>2</sup>In general this is biologically unrealistic as often the fitness effects of one mutation is correlated with those of another mutation, a situation referred to as *epistasis*.

of these focus on the actual motion of a particle along accessible paths, rather than on the paths themselves, a process known as an *adaptive walk*. A simulation study of the effect of periodically fluctuating fitness values with a random initial condition on an adaptive walk in a correlated fitness landscape was investigated in [37]. It was shown that as the fraction of sites with static fitness was increased, the dynamics of the walk changes from approximately a random walk on the entire landscape to a random walk confined to regions of high fitness. In [1], the authors also look at dynamic fitness landscapes with Gaussian correlated fitness increments and time correlations that decay exponentially with time. They use analytical and simulation methods to show that the distribution of fitness effects encountered by the walk approaches a statistical steady state. If the timescale of decorrelation is fast enough, the adaptive walk quickly ‘forgets’ its past values, whereas if the decorrelation is slower, the walk rapidly ascends to a high fitness before getting stuck and waiting for the landscape to change. To investigate the effects of an external parameter (such as an antibiotic concentration) driving the changes in the environment, the authors of [15] associate to each genotype  $\sigma$  a fitness curve  $f_\sigma(x)$  that specifies the fitness of genotype  $\sigma$  when the parameter has value  $x$ . Drawing parallels with the theory of disordered physical systems, the system is shown to display various properties such as hysteresis loops and memory formation.

### 1.1.2 Accessibility Percolation on Regular Trees

We can define accessibility percolation as follows: on an arbitrary finite rooted graph  $G = (V, E)$ , fix a subset of vertices  $T \subseteq V$  (the ‘target set’<sup>3</sup>) such that the root  $r$  of  $G$  is not in  $T$  and let  $(\Omega, \Sigma, \mathbb{P})$  be a probability space. Now to each  $v \in V$  associate a real-valued  $\Sigma$ -measurable random variable  $X(v)$  (in some contexts random variables will only be assigned to  $v \in V \setminus T$ , the idea being that the target vertices are assumed to have ‘maximal fitness’). Define a *path* in  $G$  to be a sequence of adjacent vertices  $v_0, v_1, \dots, v_h$  such that  $v_0 = r$  and  $v_h \in T$ . A path is *accessible* if the sequence of associated random variables is increasing:  $X(v_0) < X(v_1) < \dots < X(v_h)$ . In accessibility percolation we are interested in whether there exists an accessible path in  $G$  with probability 1. An alternative definition which we use below is more general and defines a path to be accessible if  $X(v_1) < X(v_2) < \dots < X(v_h)$  (i.e. a path is considered accessible if the path beginning at the neighbour of the root is increasing). The results we are concerned with in this thesis relate to the case where the fitness values are independent and uniformly distributed on  $[0, 1]$ , known in the literature as

---

<sup>3</sup>Typically  $T$  will be chosen to be the set of vertices at furthest graph-distance from the root, specifically  $T = \left\{ v \in V : d(v, r) = \max_{w \in V} d(w, r) \right\}$ , which in the case of the hypercube is a singleton.

the *House of Cards* model:

**Definition 1.1.1.** Let  $G = (V, E)$  be a graph of bounded degree. Suppose we associate to each vertex  $v \in V$  an independent random variable  $X_v \sim \text{Unif}[0, 1]$ . The **House of Cards model** on  $G$  is defined as the pair  $(G, \{X_v\}_{v \in V})$ .

In their paper [34], Roberts and Zhao consider the House of Cards model on a regular  $n$ -ary tree<sup>4</sup> and show that when the height  $h$  of the tree depends linearly on  $n$ , the ratio  $n/h$  determines whether an accessible path exists with high probability (in the limit of increasing height). The model setup is as follows: consider an  $n$ -ary tree of height  $h$ , where  $n = \lfloor \alpha h \rfloor$  for a scaling constant  $\alpha \in (0, 1)$ , and denote this graph by  $G = G(n, h)$ . Let  $P$  be the set of paths from root to a leaf in the tree, so  $|P| = n^h$ , and each path is of length  $h$ . The result from [34] that we are mainly interested in is the following:

**Theorem 1.1.2** (Theorem 1 from [34]). Let  $A_h$  denote the event that an accessible path exists in  $G(\lfloor \alpha h \rfloor, h)$  under the House of Cards model. Then we have

$$\lim_{h \rightarrow \infty} \mathbb{P}(A_h) = \begin{cases} 0 & \text{if } \alpha \leq \frac{1}{e} \\ 1 & \text{if } \alpha > \frac{1}{e}. \end{cases}$$

In [11], Chen studies a finer regime near the ‘macroscopic critical case’  $\alpha = 1/e$  in which the height  $h$  of the tree scales as  $h = en - \beta \log n$ . There it is shown that the model in fact exhibits a phase transition depending on  $\beta$ , with critical value at  $\beta = 3/2$ . In particular, in [11] the following result is proved:

**Theorem 1.1.3** (Theorem 1 from [11]). Let  $A_n$  denote the event that an accessible path exists in  $G(n, en - \beta \log n)$  under the House of Cards model. Then we have

$$\lim_{n \rightarrow \infty} \mathbb{P}(A_n) = \begin{cases} 0 & \text{if } \beta < \frac{3}{2} \\ 1 & \text{if } \beta > \frac{3}{2}. \end{cases}$$

The aim is to obtain a dynamical analogue of Theorem 1.1.3. Towards this end, we first introduce a dynamic version of the House of Cards model on a general graph  $G$ :

**Definition 1.1.4.** Let  $G = (V, E)$  be a graph of bounded degree. The **dynamic House of Cards model of rate  $\lambda$**  on  $G$  is defined as follows: associate to each vertex  $v \in V$  a continuous-time stochastic process  $(X_v(t))_t$  representing the fitness of  $v$ . This process is such that for all  $v \in V$ ,  $X_v(0) \sim \text{Unif}[0, 1]$  and independently for all  $v \in V$ , at rate  $\lambda \in (0, \infty)$  the fitness at vertex  $v$  is resampled from a  $\text{Unif}[0, 1]$  distribution.

---

<sup>4</sup>A tree in which every non-leaf vertex has  $n$  child vertices when the tree is viewed as a directed graph with the direction away from the root.

We aim to apply this dynamic process to the tree considered in [1, 5]. Consider the dynamic House of Cards model of rate 1 on the vertices of the  $n$ -ary tree and for a path  $u \in P$  define the vector-valued process  $(X_u(t))_t$  by  $X_u(t) = (X_{u(1)}(t), X_{u(2)}(t), \dots, X_{u(h)}(t))$  where  $(X_{u(i)}(t))_t$  is the stochastic process giving the fitness of the  $i$ th vertex  $u(i)$  of the path  $u$  through time. Note that a path  $u \in P$  is increasing at time  $t \in [0, \infty)$  if and only if  $X_u(t)$  belongs to the set

$$I = \left\{ (x_1, x_2, \dots, x_h) \in [0, 1]^h : x_1 < x_2 < \dots < x_h \right\}.$$

We are interested in how the limit as  $h \rightarrow \infty$  of the probability that an increasing path appears in the time interval  $[0, 1]$  depends on the value of  $\alpha$  and  $\beta$  for each of the scaling regimes in question. This question is inspired by the dynamical percolation literature, in which the problem of study is a version of Bernoulli percolation on graphs. The general formulation, first introduced in [32], is as follows: for an infinite graph  $G$  at the time 0, the configuration of edges is generated randomly by making each edge (independently) open with probability  $p \in (0, 1)$  and closed with probability  $1 - p$ . Then to each edge  $e \in E(G)$  one associates an independent rate 1 Poisson process, and at the arrival times of this process the state of edge is resampled from a Bernoulli( $p$ ) distribution (sampling a 1 indicates the edge becomes open). Without this resampling mechanism, it is known that there exists a critical parameter  $p_c(G) \in (0, 1)$  such that for  $p > p_c(G)$  there exists an infinite connected component of open edges almost surely (in which case the event of *percolation* is said to occur), while for  $p < p_c(G)$ , no such infinite component exists almost surely. For many graphs  $G$ , the value of  $p_c(G)$  is known and in addition the percolation behaviour for  $p = p_c(G)$  is known. One example is where  $G$  is taken to be the triangular lattice - in this case it is known that percolation does not occur at  $p = p_c(G)$ . However for dynamical site<sup>5</sup> percolation on the triangular lattice, Schramm and Steif [35] were able to prove the following:

**Theorem 1.1.5** (Theorem 1.3 from [35]). *Let  $G$  be the triangular lattice and consider dynamical percolation on the vertices of  $G$  with parameter  $p = p_c(G)$ . Define the subset  $\Lambda \subseteq [0, 1]$  such that  $t \in \Lambda$  if and only if at time  $t$  there exists an infinite connected component of open vertices in  $G$ . Then  $\mathbb{P}(\Lambda \neq \emptyset) = 1$ .*

What this means is that although for each fixed time, no infinite connected component of open vertices exists almost surely, one can find times  $t \in [0, 1]$  such that an infinite open cluster exists at time  $t$ . Another noteworthy result is that dynamical (edge) percolation on the integer lattice  $\mathbb{Z}^2$  exhibits exceptional times of percolation at the critical parameter  $p_c(\mathbb{Z}^2) = \frac{1}{2}$  (see [20]). A detailed overview of Theorem 1.1.5 along

---

<sup>5</sup>In *site percolation* one takes vertices (as opposed to edges) to be open or closed.

with many other results on dynamical percolation can be found in the survey of Steif [36]. The key question for our model is whether by introducing a resampling mechanism, the accessibility behaviour changes from that in the static model, in particular in the critical cases  $\alpha = 1/e$  and  $\beta = 3/2$ . Note that at the critical parameter  $\alpha = 1/e$  it is the case that any times at which an increasing path exists for a time-interval  $[0, T(h)]$  where  $T(h) \ll h^{\frac{3}{2}}$  will be ‘exceptional’ in the following sense. Let  $X(t)$  denote the state of the entire graph at time  $t$  (this takes values in  $E = [0, 1]^m$  where  $m$  is the number of vertices in the tree) and let  $I \subseteq E$  be the subset consisting of all configurations for which an accessible path exists. Let  $S \subseteq [0, T(h)]$  be the (random) set of times at which an accessible path exists and let  $\lambda(S)$  be its Lebesgue measure. Then by joint measurability of the process  $X(t)$  we can apply Fubini’s theorem

$$\mathbb{E}(\lambda(S)) = \mathbb{E}\left(\int_0^{T(h)} \mathbb{1}_{X(t) \in I} dt\right) = \int_0^{T(h)} \mathbb{P}(X(t) \in I) dt = \mathbb{P}(X(0) \in I)T(h).$$

Viewing the probability  $\mathbb{P}(X(t) \in I)$  as a function of  $h$ , Chen [11] has shown that for all  $\varepsilon > 0$ ,  $\mathbb{P}(X(0) \in I) \leq h^{-\frac{3}{2}+\varepsilon}$  when  $\alpha = 1/e$ . Therefore, given that for some  $\varepsilon > 0$ ,

$$\frac{T(h)}{h^{\frac{3}{2}-\varepsilon}} \rightarrow 0$$

as  $h \rightarrow \infty$ , one has

$$\lim_{h \rightarrow \infty} \mathbb{E}(\lambda(S)) = 0.$$

Since  $\lambda(S) \geq 0$  it follows that  $\lambda(S)$  converges to zero in probability. Hence at the critical value, any times at which an increasing path exists become ‘extremely rare’ as  $h \rightarrow \infty$  (if they exist at all). This is analogous to a similar phenomenon in the dynamical percolation literature and so we refer in the remainder of the report to the existence/non-existence of exceptional times. Note that in this thesis we are only able to prove the existence of times of accessibility percolation at the critical parameter for a time-interval  $[0, \kappa_h]$  such that  $\kappa_h \gg h^{\frac{3}{2}}$  as  $h \rightarrow \infty$ . Thus, for these times we do not have the ‘exceptionality’ property in the sense of Lebesgue measure. A similar question of exceptional times was explored for percolation on the Erdős-Rényi random graph by Roberts and Şengül in [33]. In their paper, a second moment approach is taken to prove the existence of exceptional times of a large connected component, and we take the same approach here. In addition, we adopt a similar ‘noise-sensitivity’ approach to obtain the necessary bounds on the relevant second (and mixed) moments. In general, this work can be thought of as trying to apply many of the ideas of [33] to a dynamical version of the model studied in [34].

A key tool used to prove exceptional time results in [34] is the Fourier analysis of Boolean functions on the hypercube. For any measure  $\mu$  on the Borel space  $(\{0, 1\}, \mathcal{B}(\{0, 1\}))$ , there exists an orthonormal basis  $\{\chi_S\}_{S \subseteq [n]}$  of  $\{0, 1\}^n$  such that given a square-integrable function of the form  $f : \{0, 1\}^n \rightarrow \{0, 1\}$ , one has a unique decomposition of  $f$  of the form

$$f = \sum_{S \subseteq [n]} \alpha_S \chi_S,$$

for a collection of nonnegative constants  $(\alpha_S)_{S \subseteq [n]}$ . This is known as the *Fourier decomposition* of  $f$ . A key construction needed for our results in the critical case is a representation of square-integrable functions defined on general product probability spaces, sometimes known as the *Efron-Stein decomposition*. This generalises the Fourier decomposition, in the sense that if  $f : \Omega^n \rightarrow \{0, 1\}$  for some product space  $(\Omega^n, \Sigma^{\otimes n}, \mu^{\otimes n})$  is square-integrable, we have a unique decomposition

$$f = \sum_{S \subseteq [n]} f^{=S}$$

for a collection of  $\mu^{\otimes n}$ -orthogonal functions  $\{f^{=S}\}_{S \subseteq [n]}$  (we call these functions the *Efron-Stein functions*). In [34], the expectation of the sum  $\sum_{S \subseteq [n]} \alpha_S^2$  is related to the properties of processes called *randomised algorithms* defined on  $\{0, 1\}^n$ . A key step in proving our results in Chapter 2 involves showing that the expected sum of squares of the Efron-Stein functions can similarly be related to the behaviour of randomised algorithms defined on  $\Omega^n$ . This result (Lemma 2.5.7) has general applicability and should be useful for proving noise sensitivity-type results whenever the indicator function of the event of interest can be represented as a Boolean function on a product space (with possibly uncountable component subspaces).

Our main results in Chapter 2 consist of the following theorems, which study the two different parameterisations of the height of an  $n$ -ary tree. Theorem 1.1.6 refers to the case that  $n = \lfloor \alpha h \rfloor$  for some  $\alpha > 0$  as studied for the static case in [34].

**Theorem 1.1.6.** *Consider the dynamic House of Cards model of rate 1 on an  $n$ -ary tree of height  $h$  where  $n = \lfloor \alpha h \rfloor$ . Let  $\alpha_c = 1/e$ . We have three cases: if  $\alpha < \alpha_c$ , let  $(\kappa_h)_h$  be a sequence of real numbers such that  $\kappa_h \geq 1$  and  $\kappa_h e^{-\beta h} \rightarrow 0$  for some  $\beta > -2 \log(\alpha e)$  as  $h \rightarrow \infty$ . Then*

$$\lim_{h \rightarrow \infty} \mathbb{P}(\text{there exists } t \in [0, \kappa_h] \text{ and } u \in P \text{ such that } X_u(t) \in I) = 0. \quad (1.1)$$



If  $\alpha = \alpha_c$ , let  $(\kappa_h)_h$  be such that  $\frac{h^{\frac{3}{2}} \log h}{\kappa_h} \rightarrow 0$  as  $h \rightarrow \infty$ . Then

$$\lim_{h \rightarrow \infty} \mathbb{P}(\text{there exists } t \in [0, \kappa_h] \text{ for which there exists } u \in P \text{ such that } X_u(t) \in I) = 1. \quad (1.2)$$

Finally if  $\alpha > \alpha_c$ ,

$$\lim_{h \rightarrow \infty} \mathbb{P}(\text{there exists } u \in P \text{ such that } X_u(t) \in I \text{ for all } t \in [0, 1]) = 1. \quad (1.3)$$

Note the question of existence of exceptional times of accessibility in the interval  $[0, 1]$  in the critical case  $\alpha = 1/e$  remains open, however we conjecture that such times do not exist due to the rapid decay of the first moment of the occupation time. Now consider the case where the height  $h$  of the tree is given by  $h = en - \beta \log n$  for  $\beta > 0$  as studied in [11] for the static case. In view of Theorem 1.1.2, we are looking at the critical window here, and for the dynamic case we can restrict our attention to  $\beta < 3/2$ . We have the following:

**Theorem 1.1.7.** *Consider the dynamic House of Cards model of rate 1 on an  $n$ -ary tree of height  $h$  where  $h = en - \beta \log n$ . We have two cases: first let  $\beta < 0$  and for any  $\delta > 0$  let  $(\kappa_n)_n$  be a sequence of real numbers such that  $\kappa_n n^{\beta+2+\delta} \rightarrow 0$  as  $n \rightarrow \infty$ . Then we have*

$$\lim_{n \rightarrow \infty} \mathbb{P}(\text{there exists } t \in [0, \kappa_n] \text{ and } u \in P \text{ such that } X_u(t) \in I) = 0. \quad (1.4)$$

Now let  $\beta \in [0, 3/2)$ . Then for any  $\delta > 0$  and any sequence  $(\kappa_n)_n$  such that  $\kappa_n n^{\beta - \frac{3}{2} + \delta} \rightarrow 0$  as  $n \rightarrow \infty$  we have

$$\lim_{n \rightarrow \infty} \mathbb{P}(\text{there exists } t \in [0, \kappa_n] \text{ and } u \in P \text{ such that } X_u(t) \in I) = 1. \quad (1.5)$$

Note that in part (ii) of Theorem 1.1.7, the  $\beta = 0$  case is covered by part (ii) of Theorem 1.1.6.

An interesting possibility to extend our results would be to consider a dynamic version of the model of [16], in which the static model of accessibility percolation on an  $n$ -ary tree of height  $h$  is extended. In [16], instead of counting just increasing paths, the authors count paths such that any consecutive sequence of  $k$  vertices contain at least one vertex of an increasing sub-path. They show that the probability that such a path exists as  $h \rightarrow \infty$ , tends to 1 when  $n$  grows faster than  $(h/(ek))^{\frac{1}{k}}$ . Biologically this setup can be motivated by the notion of stochastic tunnelling. For this to occur, the

mutation rate must be large enough and selection weak enough (relative to the population size) for a small sub-population of less fit individuals to exist for long enough that a new even fitter genotype can be reached which subsequently fixes in the population, allowing a fitness valley to be crossed.

### 1.1.3 Accessibility Percolation on Infinite Trees

In Chapter 3, we consider the less studied question of the existence of accessible paths on infinite trees<sup>6</sup>, first studied in [12]. This is joint work with four other collaborators: U. De Ambroggio, D. Kious, M. Ortgiese and M. Roberts. We attach i.i.d. uniform  $[0, 1]$  random variables (fitnesses) to each vertex as before and if  $\mathcal{T}$  denotes an infinite tree, let  $\partial\mathcal{T}$  denote the set of paths. For a path  $u = (o, u(1), u(2), \dots) \in \partial\mathcal{T}$  and  $i \in \mathbb{N}$ , let  $X_{u(i)}$  denote the fitness of the  $i$ th vertex and define  $X(u) = (X_{u(1)}, X_{u(2)}, \dots)$ . Here, a path  $u \in \partial\mathcal{T}$  is accessible if  $X(u)$  belongs to the set

$$I_\infty = \left\{ (x_1, x_2, \dots) \in [0, 1]^{\mathbb{N}} : x_i < x_{i+1} \forall i \in \mathbb{N} \right\}.$$

We denote the event of the existence of such a path by  $E_\infty$ , and if  $E_h$  denotes the event of existence of an accessible subpath to level  $h$ ,  $E_{h+1} \subseteq E_h$  implies  $\mathbb{P}(E_\infty) = \lim_{h \rightarrow \infty} \mathbb{P}(E_h)$ . Note that the trees considered are locally finite, so each vertex  $v$  has a finite number of descendant vertices at the next level. The probability that none of the descendant vertices at the next level have higher fitness than  $v$  is a finite product of non-zero values, and hence is non-zero. Existence of accessible paths can thus no longer occur with probability 1; the task here is to determine which trees have accessible paths with positive probability. The main existing results come from the paper [12] in which the authors consider trees such that all vertices at the same level have the same degree. Such trees are called *spherically symmetric*. The results of [12] primarily focus on the case where the degree  $d(h)$  of vertices at level  $h \geq 1$  is given by  $d(h) = \lceil (h + 1)^\alpha \rceil$ . They show the following result:

**Theorem 1.1.8** (Theorem 3.1 in [12]). *Let  $d(h) = \lceil (h + 1)^\alpha \rceil$ . Then  $\mathbb{P}(E_\infty) = 0$  if  $\alpha \leq 1$  and  $\mathbb{P}(E_\infty) > 0$  if  $\alpha > 1$ .*

It is conjectured in [12] that a necessary criterion for a nonzero value of  $\mathbb{P}(E_\infty)$  on a spherically symmetric tree is that  $\sum_{h=1}^{\infty} \frac{1}{d(h)} < \infty$  (the authors prove that this condition is sufficient). It is also remarked in [12] that this criterion does not apply to the case  $d(h) = \lceil \alpha(h + 1) \rceil$ , which is the most natural generalisation to the infinite case of the

---

<sup>6</sup>We are considering locally finite trees here - note that the large parameter limits of the trees considered in Chapter 2 are such that each vertex has infinite degree.

$n$ -ary scaling considered in Chapter 1. In this direction, we show that for all  $\alpha > 0$ , when  $d(h) = \lfloor \alpha(h+1) \rfloor \vee 1$  we have  $\mathbb{P}(E_\infty) = 0$  (Proposition 3.4.1).

Another way to extend the results of [12] is to find a criterion for accessibility on general trees i.e. not necessarily spherically symmetric ones. For this purpose we introduce the *factorial branching number* which roughly speaking measures the average number of paths to level  $h$  on the  $h^h$  scale. To define this quantity we require the notion of a *cutset*, which is a set of edges  $\pi \subseteq E$  such that all paths from the root to infinity pass through  $\pi$  and such that the removal of any edge from  $\pi$  produces a path to infinity not passing through  $\pi$ . In other words, it is a minimal set of edges separating the root from infinity. The factorial branching number is defined in a similar way to the branching number of a tree, specifically:

**Definition 1.1.9.** Let  $\Pi = \Pi(\mathcal{T})$  denote the set of cutsets of a general tree  $\mathcal{T}$ . The factorial branching number of  $\mathcal{T}$  is then defined as

$$\text{br}_f(\mathcal{T}) = \sup \left\{ \lambda > 0 : \inf_{\pi \in \Pi} \sum_{e \in \pi} (|e|!)^{-\lambda} > 0 \right\}. \quad (1.6)$$

We can contrast the above definition with the *branching number* of Lyons and Peres [29] and the *branching ruin number* of Collevocchio, Huynh and Kious [13], which are defined as

$$\text{br}(\mathcal{T}) = \sup \left\{ \lambda > 0 : \inf_{\pi \in \Pi} \sum_{e \in \pi} \lambda^{-|e|} > 0 \right\} \quad (1.7)$$

and

$$\text{br}_r(\mathcal{T}) = \sup \left\{ \lambda > 0 : \inf_{\pi \in \Pi} \sum_{e \in \pi} |e|^{-\lambda} > 0 \right\}, \quad (1.8)$$

respectively. Roughly speaking these quantities measure the amount of ‘exponential branching’ and ‘polynomial branching’ respectively. Consider the following example:

**Example 1.1.10.** Let  $\mathcal{T} = (V, E)$  be a spherically symmetric tree such that each vertex at level  $h \geq 0$  has degree  $d(h)$ , where  $d(h)$  is a positive, increasing function of  $h$ . For such a tree, it can be shown (see Appendix, Section A.3) that the infimum over cutsets in (1.6) can be replaced by the infimum over the sets of edges at each level,

which implies

$$\begin{aligned}
\text{br}_f(\mathcal{T}) &= \sup \left\{ \lambda > 0 : \inf_{h \geq 1} \sum_{|e|=h} (|e|!)^{-\lambda} > 0 \right\} \\
&= \sup \left\{ \lambda > 0 : \inf_{h \geq 1} \left( \prod_{i=0}^{h-1} d(i) \right) \frac{1}{(h!)^\lambda} > 0 \right\} \\
&= \sup \left\{ \lambda > 0 : \inf_{h \geq 1} \prod_{i=0}^{h-1} \frac{d(i)}{(i+1)^\lambda} > 0 \right\}.
\end{aligned} \tag{1.9}$$

From (1.9), it can be seen that trees whose degrees scale as  $d(h) = h^a, a > 0$  have  $\text{br}_f(\mathcal{T}) = a > 0$ , however trees on the scale  $d(h) = (\log h)^a, a > 0$  have  $\text{br}_f(\mathcal{T}) = 0$ .

Similarly to the way in which the value of the branching number determines the probability of Bernoulli percolation on a general tree, Theorem 1.1.8 suggests that having  $\text{br}_f(\mathcal{T}) < 1$  or  $\text{br}_f(\mathcal{T}) > 1$  should determine the probability of accessibility on a general tree. Recall the definition of the House of Cards model from Definition 1.1.1 in Section 1.1.2. The main result of Chapter 3 is the following theorem covering locally finite infinite trees:

**Theorem 1.1.11.** *Let  $\mathcal{T}$  be a locally finite, infinite tree and consider the House of Cards model on  $\mathcal{T}$ . For  $h \geq 1$ , let  $E_h$  denote the event that there exists an accessible path to level  $h$  and define  $E_\infty = \cap_{h \geq 1} E_h$ . Then  $\mathbb{P}(E_\infty) > 0$  if  $\text{br}_f(\mathcal{T}) > 1$  and  $\mathbb{P}(E_\infty) = 0$  if  $\text{br}_f(\mathcal{T}) < 1$ .*

The proof of the above result is contained in Propositions 3.3.1 and 3.3.2. The idea is to use the theory of electrical networks on trees as detailed in the book of Lyons and Peres [29]. To do this, we first specify a sequence of positive numbers  $(p_n)_n$  such that  $\sum_{n \geq 1} p_n \leq 1$  and attach independent *Bernoulli*( $p_n$ ) random variables to each vertex at level  $n$ . If the random variable sampled from a vertex takes the value 1 then we say the vertex is *open*; thus at level  $n$  each vertex has probability  $p_n$  of being open. We choose the  $(p_n)_n$  in such a way that the event of accessibility percolation is implied by the existence of an open path to infinity.

Note that the critical case  $\text{br}_f(\mathcal{T}) = 1$  is indeterminate for general trees. For spherically symmetric trees, this critical window can be explored in more detail, with a key open problem in this area being to prove (or find a counterexample to) the following conjecture:

**Conjecture 1.1.12.** *Let  $\mathcal{T}$  be a spherically symmetric tree with degree sequence  $(d(i))_{i \geq 1}$*

satisfying

$$\sum_{i=1}^{\infty} \frac{1}{d(i)} = \infty$$

Then  $\mathbb{P}(E_{\infty}) = 0$ .

Note that in view of the results of [12] mentioned earlier, a proof of Conjecture 1.1.12 would fully characterise the accessibility of spherically symmetric trees.

## 1.2 Adaptive Dynamics

### 1.2.1 General Theory and Motivation

In Chapter 4, we move on to the consideration of a model of population dynamics and of primary interest is to discover how a periodic variation in the environment affects the evolution of populations. We will define a model that falls within the framework of *adaptive dynamics* models, first introduced in a rigorous fashion by Champagnat [8]. This class of models enables one to model the growth of multiple subpopulations of a larger population whilst including the effect of competition for resources. The amount of resources (in this case nutrients) available is encapsulated in an integer parameter  $K$  which becomes the ‘scale’ on which the size of the subpopulations are measured. Specifically, the class of models within the adaptive dynamics framework constrain the total population sizes (of all subpopulations combined) to be at most a scalar multiple of  $K$ . To study the qualitative behaviour of the subpopulations, independent of a particular choice of  $K$ , the aim will be to calculate the limiting probabilities, as  $K \rightarrow \infty$ , that a given type becomes fixed.

First we outline the well-studied case in which only one population is present, beginning with a single individual, such that each individual has birth rate  $b > 0$  and death rate  $d > 0$  with  $b > d$  (so the population grows on average). We refer to the difference  $b - d$  as the growth rate. In the case of no competition between individuals, the situation is typically modelled as a birth-death process. In a birth-death process, the rates  $b$  and  $d$  are interpreted as follows: when an individual is born at time  $t \geq 0$ , two independent random times are assigned to it - a birth time  $T_b$ , distributed exponentially with parameter  $b$  and a death time  $T_d$ , distributed exponentially with parameter  $d$ . If  $T_b < T_d$  then a new individual is born at time  $t + T_b$  and a pair of independent random times are assigned to the newly born individual, as well as a new pair of random times to the parent. If instead  $T_d < T_b$ , the individual does not give birth and dies at time  $t + T_d$ . From the properties of exponential distributions and the assumption of

independence among individuals, the events  $\{T_b < T_d\}$  and  $\{T_d < T_b\}$  can be shown to have probabilities  $b/(b+d)$  and  $d/(b+d)$  respectively. However, the above assumes no competition for resources among the individuals of the population. The adaptive dynamics framework models the effect of such competition by supposing that the death rate of each individual should not be constant but should be state-dependent: specifically in such a way that when the total population is large (which we measure relative to  $K$ ) the death rate of individuals is increased. Thus, the above setup is modified so that if  $N(t)$  denotes the population size at time  $t \geq 0$ , an individual born at time  $t$  has birth rate  $b$  as before and death rate  $d + CN(t)/K$ , where  $C > 0$  is a parameter controlling the strength of competition between individuals. This new process is referred to as a logistic birth-death process, whose properties were studied in [28]. A notable result of [28] (Theorem 2.2) is that if  $d = 0$ , i.e. deaths only occur due to competition, the process  $N(t)$  is positive recurrent in  $\mathbb{N}$  and the stationary distribution is a Poisson random variable of parameter  $\frac{b}{C}$  conditioned to be non-zero.

The growth of a logistic birth-death process beginning with one individual can be divided into (up to) three phases. To specify the phases we need to fix a small constant  $\varepsilon > 0$  and parameters  $b$ ,  $d$  and  $C$  as detailed above. The first phase lasts from time 0 until the first time the population size exceeds  $\varepsilon K$ , which we denote by  $T_{\varepsilon K}$ . In this phase the population size is small so stochastic effects dominate, and there is a possibility due to this that the population will become extinct before it ever reaches size  $\varepsilon K$ . In addition, due to the small population size the effect of competition between the populations is negligible. This means the population size can be approximated by that of a simple birth-death process with birth rate  $b$  and death rate  $d$ . On average the population will therefore grow exponentially with growth rate  $b-d$ , so solving the equation  $\varepsilon K = e^{(b-d)T_{\varepsilon K}}$  shows that, given the population does not go extinct, we expect the duration of phase 1 to be approximately equal to  $T_{\varepsilon K} = \left(\frac{1}{b-d}\right) \log K - \frac{\varepsilon}{b-d} \approx \left(\frac{1}{b-d}\right) \log K$ , where the approximation is valid for large  $K$  and small  $\varepsilon$ .

The second phase lasts from time  $T_{\varepsilon K}$  until  $\bar{T}$ , the first time the population size enters the interval  $I_\varepsilon = [K(\bar{n}-\varepsilon), K(\bar{n}+\varepsilon)]$ , where  $\bar{n} = \frac{b-d}{C}$  is the stable steady state solution of the autonomous differential equation

$$\frac{dn}{dt} = (b-d-Cn)n \quad , \quad n(0) = \varepsilon . \quad (1.10)$$

This means that since we assume  $\varepsilon > 0$  is small,  $N(\bar{T})/K \approx \bar{n}$ . Once the second phase is over, the population is said to be *resident*. During the second phase, the logistic

birth-death process governing the evolution of the population size can be approximated using the ODE (1.10). The solution of such an equation can be shown to reach the interval  $I_\varepsilon$  in a time of constant order with respect to  $K$ . This means the duration of phase 2 is very quick - it is negligible on the  $\log K$  timescale for large  $K$  compared to phase 1 which takes a time of approximately  $1/(b-d)$  on the  $\log K$  timescale. It should also be mentioned that due to the starting population size of approximately  $\varepsilon K$  at the beginning of phase 2, the probability of extinction approaches 0 in the limit of large  $K$  (in contrast to phase 1 where we start with 1 individual - in this case the probability of going extinct before reaching size  $\varepsilon K$  approaches the value  $d/b$  for large  $K$ ). The third phase begins at time  $\bar{T}$ . This phase is largely uninteresting in the absence of additional subpopulations as the population size can be shown to remain with high probability within the interval  $I_\varepsilon$  for a time of length  $e^{VK}$  for some  $V > 0$  (though it will go eventually extinct, almost surely).

The first paper proving rigorously the existence of the three phases described heuristically above, was by Champagnat [8], which studied the general case of multi-type populations. The model included a mutation mechanism of rate  $\mu_K$ , where the set of possible types is a compact subset<sup>7</sup> of  $\mathbb{R}^l$ ,  $l \geq 0$ . The results in [8] were proved under the assumption that  $\mu_K K \log K \rightarrow 0$  and  $e^{VK} \mu_K \rightarrow \infty$  for all  $V > 0$ , as  $K \rightarrow \infty$ . The theory often admits extensions to multi-type populations experiencing mutation events as well as complex interactions between individuals such as horizontal gene transfer (as in [10]). To deal with multiple populations experiencing mutation and selection, a measure-valued formulation of the theory is generally required, which was first introduced by Fournier and Méléard [19] and adapted in [8]. In our case, however we will focus on only two types and not allow mutation to other types, although we will modify the model in a different direction by allowing the birth and death rates of individuals of a particular type to depend on time. The model we study is a partially time-dependent extension of the basic model studied by Billiard and Smadi [5]. In [5], the effects of different interactions between two newly appeared mutants, during the course of their invasion of an existing population of a third type, are explored. Of interest is the probability that either of the initiated mutants succeeds in supplanting the non-mutated population as the only (or largest) subpopulation across all the types, as well as estimates of the duration of different phases of the invasion and the duration of the invasion as a whole.

---

<sup>7</sup>If the compact subsets are assumed to be closed balls centered at the origin, the setup becomes somewhat similar to Fisher's geometric model.

## 1.2.2 Invasion in a Periodic Environment

Our model consists of a resident type, which we call type 0, and an invading type, which we call type 1. We specify the resident type to have population size  $\bar{n}_0 K$  at time 0 (we will specify the value of the constant  $\bar{n}_0$  later) and the invading type begins with one individual. We suppose the intrinsic birth rate of the invading population is dependent on an external environment which is constant for periods of length  $S_K$  and fluctuates between two states. Note that the period length  $S_K$  is allowed to depend on  $K$  and different choices will lead to different probabilities of invasion for the type 1 population. More generally, instead of only allowing the birth rate to vary between two possible states, we can specify  $2S_K$ -periodic functions  $b_1(t)$  and  $d_1(t)$  which represent the birth and death intensities for the invading population, and require that both functions are nonnegative, bounded and piecewise continuous. Some of our preliminary lemmas will be proved in this more general context. Meanwhile, we suppose throughout that the resident type has constant intrinsic birth and death rates,  $b_0(t) = b_0$  and  $d_0(t) = d_0$  respectively. The process we are interested in is then a pair  $(N_0(t), N_1(t))_{t \geq 0}$  such that  $(N_0(0), N_1(0)) = (\bar{n}_0 K, 1)$  and for  $t \geq 0$ , given a birth or death event occurs at time  $t$ , the time until the next birth or death event is  $T(t) = T_b^0(t) \wedge T_d^0(t) \wedge T_b^1(t) \wedge T_d^1(t)$ , where for each  $i \in \{0, 1\}$ ,  $j \in \{0, 1\} \setminus \{i\}$ ,  $T_b^i(t)$  is distributed as

$$\mathbb{P}(T_b^i(t) \leq s) = 1 - \exp \left\{ - \int_t^{t+s} N_i(u) b_i(u) du \right\}$$

and  $T_d^i(t)$  is distributed as

$$\mathbb{P}(T_d^i(t) \leq s) = 1 - \exp \left\{ - \int_t^{t+s} N_i(u) \left( d_i(u) + \frac{C_{i,i} N_i(u)}{K} + \frac{C_{i,j} N_j(u)}{K} \right) du \right\}.$$

In the case where the function  $b_1(t)$  only takes two values,  $b_1^+$  in odd periods and  $b_1^- < b_1^+$  in even periods with constant death rate  $d_1$ , the joint process can be formulated as follows:

- For all  $t \in [0, \infty)$ , each type 0 individual gives birth at rate  $b_0$  and dies at rate  $d_0 + \frac{C_{0,0} N_0(t)}{K} + \frac{C_{0,1} N_1(t)}{K}$ .
- At times  $t \in \bigcup_{m=0}^{\infty} [2mS_K, (2m+1)S_K)$ , each type 1 individual gives birth at rate  $b_1^+$  and dies at rate  $d_1 + \frac{C_{1,1} N_1(t)}{K} + \frac{C_{0,1}(t) N_0(t)}{K}$ .
- At times  $t \in \bigcup_{m=0}^{\infty} [(2m+1)S_K, 2(m+1)S_K)$ , each type 1 individual gives birth at rate  $b_1^-$  and dies at rate  $d_1 + \frac{C_{1,1} N_1(t)}{K} + \frac{C_{0,1}(t) N_0(t)}{K}$ .



As in the time-independent framework discussed earlier, during the first phase (in which the type 1 population has size no greater than  $\varepsilon K$  for all  $\varepsilon > 0$  and the type 0 population stays close to its equilibrium size) the competition pressure exerted on the type 1 population remains approximately constant. Therefore we can model the evolution of the type 1 population size as a periodic birth-death process (without density dependence) but with an adjusted individual death rate. The general theory of birth-death processes with time-dependent rates is detailed in the paper of Kendall [24]. A key result in [24] is that the distribution of the time to extinction  $T_0$  of a birth-death process with instantaneous birth and death rates  $b(t)$  and  $d(t)$  at time  $t \geq 0$ , starting from 1 individual, is determined by

$$\mathbb{P}(T_0 \leq t) = \frac{\int_0^t e^{\rho(s)} d(s) ds}{1 + \int_0^t e^{\rho(s)} d(s) ds}, \quad (1.11)$$

where  $\rho(t) = \int_0^t (d(s) - b(s)) ds$ . The above formula identifies the key role played by the function  $\rho(t)$  in determining the long-time behaviour of the process. In the time-dependent context, the analog of a supercritical process (one with probability of eventual extinction less than 1) is a process for which  $\int_0^\infty e^{\rho(s)} d(s) ds < \infty$ . In the case that  $d(t)$  is bounded, (which is the case we consider with only two finite values taken by death rate function  $d(t)$ ) this condition reduces to the requirement that  $\int_0^\infty e^{\rho(s)} ds < \infty$ . A sufficient condition for any  $K$  is to require that  $\sup_{t \geq 0} \frac{\rho(t)}{t} < 0$ . For the biological processes formulated above we will assume in our main result (stated below) that  $d_1 < b_1^-$  so that the process is always supercritical in a given period. Our main result, Theorem 1.2.1, concerns the probability that the type 1 population reaches size  $\varepsilon K$ , which is a prerequisite for successfully invading the population. To state the theorem we need some additional notation. For  $i \in \{0, 1\}$  and  $R > 0$  introduce the stopping times

$$T_R^{(i)} = \inf \{t \geq 0 : N_i(t) \geq R\}$$

and

$$T_0^{(i)} = \inf \{t \geq 0 : N_i(t) = 0\}.$$

Also we define the relevant  $\varepsilon K$ -neighbourhood of the type 0 population as

$$I_\varepsilon^{(0)} = \left[ K \left( \bar{n}_0 - \frac{\varepsilon C_{1,0}}{C_{0,0}} \right), K \left( \bar{n}_0 + \frac{\varepsilon C_{1,0}}{C_{0,0}} \right) \right]$$

as well as the stopping time

$$\tilde{T}_\varepsilon^{(0)} = \inf \{t \geq 0 : N_0(t) \notin I_\varepsilon^{(0)}\}.$$

**Theorem 1.2.1.** *Let the period  $S_K$  be such that  $S_K/\log K \rightarrow 0$  as  $K \rightarrow \infty$ . Suppose also that  $d_1(t) = d_1 < b_1^-$ , a constant, for all  $t \geq 0$ . If  $S_K \rightarrow \infty$ , there exist constants  $R$  and  $R'$  such that for any  $\varepsilon > 0$ ,*

$$(i) \quad \mathbb{P}(T_0^{(1)} < \tilde{T}_\varepsilon^{(0)}) \rightarrow \frac{d_1}{b_1^+} - R\varepsilon \quad \text{as } K \rightarrow \infty$$

$$(ii) \quad \mathbb{P}(T_{\varepsilon K}^{(1)} < \tilde{T}_\varepsilon^{(0)}) \rightarrow \left(1 - \frac{d_1}{b_1^+}\right) - R'\varepsilon \quad \text{as } K \rightarrow \infty$$

*If instead  $S_K \rightarrow 0$  as  $K \rightarrow \infty$ , there exists constants  $R$  and  $R'$  such that for any  $\varepsilon > 0$ ,*

$$(i) \quad \mathbb{P}(T_0^{(1)} < \tilde{T}_\varepsilon^{(0)}) \rightarrow \frac{d_1}{\frac{1}{2}(b_1^+ + b_1^-)} - R\varepsilon \quad \text{as } K \rightarrow \infty$$

$$(ii) \quad \mathbb{P}(T_{\varepsilon K}^{(1)} < \tilde{T}_\varepsilon^{(0)}) \rightarrow \left(1 - \frac{d_1}{\frac{1}{2}(b_1^+ + b_1^-)}\right) - R'\varepsilon \quad \text{as } K \rightarrow \infty$$

*Moreover, for sufficiently large  $K$  there exists  $\beta > 0$  such that*

$$\mathbb{P}\left((1 - c\varepsilon)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1} < T_{\varepsilon K}^{(1)} < (1 + c\varepsilon)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1} \mid T_{\varepsilon K}^{(1)} < \tilde{T}_\varepsilon^{(0)}\right) \geq 1 - \beta\varepsilon.$$

As the period length depends on  $K$ , we will get different behaviour for different scalings of  $S_K$ : if  $S_K \rightarrow \infty$  while  $S_K/\log K \rightarrow 0$ , the probability the invading population reaches size  $\varepsilon K$  is found to depend only on the first period, as this period is ‘long enough’ in the limit to guarantee whether the type 1 population will survive or go extinct before establishing itself. However if  $S_K \rightarrow 0$ , the probability of reaching size  $\varepsilon K$  will reflect an averaging between the birth and death rates of the odd and even periods. The approximate duration of the time to reach size  $\varepsilon K$ , given this occurs, is however a function of the average growth rate,  $\frac{1}{2}(b_1^- + b_1^+) - d_1$ .

In Section 4.4, we consider the behaviour of the two populations once the type 1 population has reached size  $\varepsilon K$ . In the case where  $S_K \rightarrow \infty$ , suppose the type 1 population first reaches size  $\varepsilon K$  in the  $i_K$ th period for some  $i_K \in \mathbb{N}$ , that is,  $T_{\varepsilon K}^{(1)} \in [(i_K - 1)S_K, i_K S_K)$ . Let  $b_{i_K}^1$  denote the birth rate of the type 1 population during  $[(i_K - 1)S_K, i_K S_K)$  and define  $y_0 = \frac{N_0(T_{\varepsilon K}^{(1)})}{K}$ . Then as  $K \rightarrow \infty$ , the pair of population density processes given by  $\mathbf{n}(t) = \left(\frac{N_0(t - T_{\varepsilon K}^{(1)})}{K}, \frac{N_1(t - T_{\varepsilon K}^{(1)})}{K}\right)$  will approximate

the solution  $(n_0(t), n_1(t))$  of the system of ODEs

$$\begin{aligned}\frac{dn_0}{dt} &= (b_0 - d_0 - C_{0,0}n_0 - C_{0,1}n_1)n_0 \quad , \quad n_0(0) = y_0, \\ \frac{dn_1}{dt} &= (b_{i_K}^1 - d_1 - C_{0,1}n_0 - C_{1,1}n_1)n_1 \quad , \quad n_1(0) = \varepsilon,\end{aligned}\tag{1.12}$$

uniformly on any finite interval  $[0, T]$ ,  $T \leq i_K S_K - T_{\varepsilon K}^{(1)}$ , with probability tending to 1 as  $K \rightarrow \infty$ . This follows from existing results, e.g. Lemma A.2 of [5]. As soon as the period changes, the system (1.12) with  $b_{i_K}$  replaced by  $b_{i_{K+1}}$  will then determine the dynamics, and the pair  $\mathbf{n}(t)$  will in fact reach a neighbourhood of a deterministic steady state within the  $(i_K + 1)$ th period, since the system (1.12) with  $b_{i_K}$  replaced by  $b_{i_{K+1}}$  reaches its stable steady state in  $O(1)$  time, while  $S_K \rightarrow \infty$ . If instead  $S_K \rightarrow 0$ , we will show in Proposition 4.4.1 that for  $O(1)$  time, the pair of population density processes stays close to an ‘averaged’ version of (1.12) with probability tending to 1 as  $K \rightarrow \infty$ .

### 1.2.3 Invasion with Delay

One question that can be asked in relation to the competition between two mutants to become the resident type in a population, is how does a delay in the introduction of one of the mutants affect the probability of fixation? This question was inspired by the paper [21] in which the mutational routes to recover motility in initially immotile variants of a particular bacterial strain were studied. In the main experiment, a homogeneous population of the bacterium *Pseudomonas fluorescens* with motility genes removed was inoculated into the centre of an agar plate. The bacteria are able to consume nutrients from the surrounding medium however require motility to be able to consume nutrients further away from the centre. At some time after the first motile mutant emerges, typically 2-3 days, the leading edge of the bacterial growth frontier was sampled to determine which mutant made up the majority of the surviving population. It was found that the evolution was strongly biased towards one of the possible mutational routes to a motile phenotype. One particular mutation was found to occur much more frequently than the others - this is the mutation A289C, known as the ‘hotspot mutation’. A possible explanation for this bias is a difference in fitness effects, which would be the case if the hotspot mutation conferred faster motility than the other mutants. By comparing the A289C mutant to a less commonly occurring mutant, A683C, this was found not to be the case. Both mutants are the result of a single nucleotide change (an A to C transversion) and are located in the same region of the bacterial genome. The two mutants were grown separately across

four environmental conditions for 48 hours. The differences in the migration area in each environment were found to be negligible. In another experiment from the same paper, equal populations of the A289C and A683C mutants were inoculated at the same time onto a single plate and allowed to compete with each other. The bacterial growth frontier was sampled after 24 hours. This experiment was repeated, and in 1 out of 4 independent experiments the cells sampled were found to be of type A683C (the other 3 experiments found A289C). This is a much higher proportion than was found in the original evolution experiment starting from an immotile population, in which only 1 out of 116 independent runs ended with A683C being the dominant type (although it should be noted that the difference in the number of replicates of the two experiments means the 1/4 estimate has a lower statistical confidence<sup>8</sup>). In addition, the effect of early appearance of one of the types on the outcome of evolution was tested by inoculating the A289C and A683C mutants onto the same plate at different times. It was observed that after a delay of 6 or more hours in the introduction of the A289C cells, the A683C will always dominate the population. These results suggest that the bias is caused by the hotspot mutation appearing more rapidly than the others in the population and it is this ‘early appearance’ advantage rather than higher fitness that allows it to dominate the population. In Chapter 5, the purpose will be to investigate mathematically this question of fixation with delay, in particular, if we introduce a noticeable fitness difference (unlike in the experiments detailed above), and allow a less fit mutant to begin growing before the introduction of a fitter type, how large must the delay be for the less fit mutant to nevertheless become the resident type?

To simplify the modelling of the situation, we can interpret the fitness differences via a difference in growth rates of the two types, due to the fact that higher motility leads to a higher birth rate. We again suppose that both populations interact and evolve according to coupled logistic birth-death processes in a similar way to Chapter 4 (but with time-independent birth and death rates). We find that, in a particular sense, the less fit type in a pair of mutants can achieve fixation: if  $b_0, d_0$  and  $b_1, d_1$  denote the birth and death rates of types 0 and 1 respectively, with  $0 < b_0 - d_0 < b_1 - d_1$ , then a large enough delay (dependent on the initial population sizes of the two populations) ensures that the type 0 population reaches size  $\varepsilon K$  for some small  $\varepsilon > 0$  before type 1. This will not on its own ensure fixation, however from existing theory if the *invasion*

---

<sup>8</sup>We can represent the original experiment as a sequence of 116 Bernoulli trials with true parameter  $p_1$  and the competition experiment with equal initial populations as a sequence of 4 Bernoulli trials with true parameter  $p_2$ . Let  $X_1$  and  $X_2$  denote the sums of the outcomes of the first and second experiments respectively, and define the estimators  $\hat{p}_1 = X_1/116$  and  $\hat{p}_2 = X_2/4$ . Then by Chebyshev’s inequality, the probability that  $\hat{p}_1$  differs by more than  $\varepsilon > 0$  from  $p_1$  is no more than  $p_1(1 - p_1)/116\varepsilon^2$  whereas the probability that  $\hat{p}_2$  differs from  $p_2$  by more than  $\varepsilon$  is no more than  $p_2(1 - p_2)/4\varepsilon^2$ .

fitness  $F_{1,0}$  of the type 1 population with respect to the type 0, defined by<sup>9</sup>

$$F_{1,0} = b_1 - d_1 - C_{0,1}\bar{n}_0 = b_1 - d_1 - \frac{C_{0,1}}{C_{0,0}}(b_0 - d_0),$$

is negative, then the type 0 population will successfully fixate. This is only possible if the competition between populations is greater than the competition within populations. The timescale for a birth-death process to reach polynomial size in  $K$  is of order  $\log K$  for large  $K$ , and so we suppose that the type 1 population is introduced after a time  $\alpha \log K$  for some  $\alpha > 0$ . The main result of this chapter, Theorem 1.2.2, determines a value  $\alpha_* > 0$  (again dependent on the initial conditions of the two populations) such that for  $\alpha < \alpha_*$  the probability of invasion approaches a value  $p_-$  as  $K \rightarrow \infty$  and for  $\alpha > \alpha_*$  the probability of invasion approaches a value  $p_+$ , where  $0 \leq p_- < p_+ \leq 1$ . The starting sizes of the two populations we take as  $K^{\gamma_0}$  and  $K^{\gamma_1}$  with  $\gamma_0, \gamma_1 \in [0, 1)$ . As in the previous section, for each  $i \in \{0, 1\}$  and  $R > 0$  we define the stopping times

$$T_R^{(i)} = \inf \{t \geq 0 : N_i(t) \geq R\}$$

and

$$T_0^{(i)} = \inf \{t \geq 0 : N_i(t) = 0\}.$$

We also define the time delay  $T_D = \alpha \log K$ .

**Theorem 1.2.2.** *Suppose  $0 < b_0 - d_0 < b_1 - d_1$  and let  $\gamma_0, \gamma_1 \in [0, 1)$  be such that*

$$(\gamma_1 - 1) > (\gamma_0 - 1) \left( \frac{b_1 - d_1}{b_0 - d_0} \right). \quad (1.13)$$

*If  $0 < \alpha < \frac{1-\gamma_0}{b_0-d_0} - \frac{1-\gamma_1}{b_1-d_1}$  there exists  $R > 0$  such that for  $\varepsilon > 0$  sufficiently small, we have as  $K \rightarrow \infty$ ,*

- *if  $\gamma_0 \geq 0, \gamma_1 > 0$ ,*

$$\mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}) \rightarrow 1 \quad \text{and} \quad \mathbb{P}(T_{\varepsilon K}^{(0)} < T_{\varepsilon K}^{(1)}) \rightarrow 0$$

- *if  $\gamma_0 > 0, \gamma_1 = 0$ ,*

$$\left| \mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}) - \left(1 - \frac{d_1}{b_1}\right) \right| < R\varepsilon \quad \text{and} \quad \left| \mathbb{P}(T_{\varepsilon K}^{(0)} < T_{\varepsilon K}^{(1)}) - \frac{d_1}{b_1} \right| < R\varepsilon$$

---

<sup>9</sup>Recall that  $\bar{n}_0$  denotes the deterministic nonzero steady state of the type 0 population density according to the ODE (1.10).

- if  $\gamma_0 = 0, \gamma_1 = 0$ ,

$$\left| \mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}) - \left(1 - \frac{d_1}{b_1}\right) \right| < R\varepsilon \quad , \quad \left| \mathbb{P}(T_{\varepsilon K}^{(0)} < T_{\varepsilon K}^{(1)}) - \left(1 - \frac{d_0}{b_0}\right) \frac{d_1}{b_1} \right| < R\varepsilon$$

and  $\left| \mathbb{P}(T_0^{(0)} \vee T_0^{(1)} < T_{\varepsilon K}^{(0)} \wedge T_{\varepsilon K}^{(1)}) - \frac{d_0 d_1}{b_0 b_1} \right| < R\varepsilon$

If instead  $\alpha > \frac{1-\gamma_0}{b_0-d_0} - \frac{1-\gamma_1}{b_1-d_1}$  there exists  $R > 0$  such that for  $\varepsilon > 0$  sufficiently small we have as  $K \rightarrow \infty$ ,

- if  $\gamma_0 > 0, \gamma_1 \geq 0$ ,

$$\mathbb{P}(T_{\varepsilon K}^{(0)} < T_{\varepsilon K}^{(1)}) \rightarrow 1 \quad \text{and} \quad \mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}) \rightarrow 0$$

- if  $\gamma_0 = 0, \gamma_1 > 0$

$$\left| \mathbb{P}(T_{\varepsilon K}^{(0)} < T_{\varepsilon K}^{(1)}) - \left(1 - \frac{d_0}{b_0}\right) \right| < R\varepsilon \quad \text{and} \quad \left| \mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}) - \frac{d_0}{b_0} \right| < R\varepsilon$$

- if  $\gamma_0 = 0, \gamma_1 = 0$ ,

$$\left| \mathbb{P}(T_{\varepsilon K}^{(0)} < T_{\varepsilon K}^{(1)}) - \left(1 - \frac{d_0}{b_0}\right) \right| < R\varepsilon \quad , \quad \left| \mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}) - \frac{d_0}{b_0} \left(1 - \frac{d_1}{b_1}\right) \right| < R\varepsilon$$

and  $\left| \mathbb{P}(T_0^{(0)} \vee T_0^{(1)} < T_{\varepsilon K}^{(0)} \wedge T_{\varepsilon K}^{(1)}) - \frac{d_0 d_1}{b_0 b_1} \right| < R\varepsilon$

In particular, if  $\alpha > \frac{1-\gamma_0}{b_0-d_0}$  and  $\gamma_0 > 0$ ,

$$\mathbb{P}(T_{\varepsilon K}^{(0)} < T_D) \rightarrow 1$$

as  $K \rightarrow \infty$ .

## Chapter 2

# Dynamic Accessibility Percolation on Regular Trees

### 2.1 Introduction

The aim of this chapter is to prove Theorems 1.1.6 and 1.1.7, both of which relate to dynamic accessibility percolation on  $n$ -ary trees<sup>1</sup> of height  $h$ , where  $n$  and  $h$  are related through a given parameter. We consider two parameterisations: Theorem 1.1.6 looks at the case where the  $n = \alpha h$  for some  $\alpha > 0$  and Theorem 1.1.7 looks at the case where  $h = en - \beta \log n$  for some  $\beta \in \mathbb{R}$ . For the first parameterisation we consider three parameter regions:  $\alpha < 1/e$  (subcritical),  $\alpha = 1/e$  (critical) and  $\alpha > 1/e$  (supercritical). For the second parameterisation we consider the parameter regions  $\beta < 0$  (strongly subcritical), and  $\beta \in [0, 3/2)$  (critical window).

We recall the general setup and notation introduced in Section 1.1.2. We consider the dynamic House of Cards model of rate 1 (see Definition 1.1.4 in Section 1.1.2) on the vertices of an  $n$ -ary tree with vertex set  $V$ , and for each vertex  $v \in V$  we let  $X_v(t)$  denote the fitness of  $v$  at time  $t \geq 0$ . Let  $P$  denote the set of paths in the tree and for  $u \in P$ ,  $i \in \{1, 2, \dots, h\}$  let  $u(i)$  denote the  $i$ th vertex from the root along path  $u$ . Now define  $X_u(t) = (X_{u(1)}(t), X_{u(2)}(t), \dots, X_{u(h)}(t)) \in [0, 1]^h$  and the set

$$I = \left\{ (x_1, x_2, \dots, x_h) \in [0, 1]^h : x_1 < x_2 < \dots < x_h \right\}.$$

---

<sup>1</sup>Recall that an  $n$ -ary tree is tree such that each vertex has  $n$  descendant vertices at the next level.

Using this notation, the event that a path  $u \in P$  is accessible at time  $t \geq 0$  can be written as  $\{X_u(t) \in I\}$ . It will also be useful for later on to introduce, for  $1 \leq K \leq h$  the set

$$I_K = \left\{ (x_1, x_2, \dots, x_K) \in [0, 1]^K : x_1 < x_2 < \dots < x_K \right\}.$$

Additionally, we will make extensive use of *Stirling's approximation* which we state in the following simplified form:

**Fact 2.1.1** (Stirling's Approximation). *For all positive integers  $n \geq 1$  we have*

$$2e^{-n}n^{n+\frac{1}{2}} < n! < 3e^{-n}n^{n+\frac{1}{2}}. \quad (2.1)$$

For a proof see Theorem 2.1 in [4], which proves a more detailed result<sup>2</sup> of which the above statement is a straightforward corollary.

## 2.2 Overview of Proofs

The proofs in this chapter are structured as follows: in Section 2.2 we concentrate on the subcritical case of the first parameterisation and provide the proof of Proposition 2.3.1 as well as part 1 of Theorem 1.1.6 which follows as a simple corollary of this result. We then prove the subcritical result for the second parameterisation which is part 1 of Theorem 1.1.7. In Section 2.3 we turn to the supercritical case for the first parameterisation and prove part 3 of Theorem 1.1.6. In Section 2.4 we provide the proof of part 2 of Theorem 1.1.6 which looks at the critical case for the first parameterisation (in fact we show the more detailed result Proposition 2.5.1). This is made up of a number of components. Let  $P$  denote the set of paths (from the root to a terminal vertex) in the tree and for paths  $u \in P$ , we define the random variable  $Z_u$  to be the total time in  $[0, \kappa_n]$  during which the vector of fitness values of the path  $u$  falls in a particular subset  $J_n \subseteq [0, 1]^h$  (where  $J_n$  is chosen such that a path with fitness values in  $J_n$  is accessible). The key to the proof is to obtain adequate bounds on expectations of the form  $\mathbb{E}(Z_u Z_v)$  for random variables  $Z_u$  and  $Z_v$  corresponding to pairs of paths  $(u, v)$  in the tree. In particular, we introduce a modified setup as in [34] and [11] by choosing  $J_n$  to be a strict subset of the set of possible accessible fitness vectors, and hence we only count a subset of the set of accessible paths in the tree at any given time  $t$ . This allows better control of the second and mixed moments. We look firstly at the case of pairs of independent paths and the case where  $u = v$ . Then for paths that agree up to an intermediate point, we look at the correlations over 'short' and 'long' time intervals,

---

<sup>2</sup>The result of [4] shows that the constant 3 can be replaced by  $2\sqrt{\frac{6\pi}{11}}$  in the right-hand side of (2.1).



which are covered in Sections 2.4.2 and 2.4.3 respectively. At the end of Section 2.4 the preceding calculations are combined to finish the proof of Proposition 2.5.1 which implies part 2 of Theorem 1.1.6. Finally in Section 2.5 we prove part 2 of Theorem 1.1.7 which requires us to look at the time interval needed to see an accessible path for the subcritical to critical case of the second parameterisation, that is  $\beta \in (0, 3/2)$ . This involves a “boosting” argument in which we use the results of [11] to show that there are many accessible and non-resampled paths to level  $K = \gamma \log n$  in the tree for  $\gamma > 0$  depending on  $\beta$ . The idea is that the probability that none of these paths possess an extension to level  $h$  which has exceptional times making the whole path accessible is small. This is formalised using a second moment method. Throughout the remaining sections, we order the set of paths  $P$  according to the position of the terminal vertex of each element, starting from the left (when the tree is embedded in the plane). Thus the path with the left-most terminal vertex will be given the number 1.

## 2.3 Subcritical Cases

We first prove part (i) of Theorem 1.1.6 which follows as a corollary of the following Proposition:

**Proposition 2.3.1.** *Let  $\alpha_c = 1/e$ . Then if  $\alpha < \alpha_c$  there exists a  $\beta > 0$  such that for large enough  $h$ ,*

$$\mathbb{P}(\text{there exists } t \in [0, 1] \text{ for which there exists } u \in P \text{ such that } X_u(t) \in I) \leq e^{-\beta h}. \quad (2.2)$$

*Proof of Proposition 2.3.1.* For a Borel subset  $B \in \mathcal{B}([0, 1])$  and a path  $u \in P$ , let  $A_{u,B}$  be the event that  $u$  is increasing at some time in  $B$  and let  $A_B$  be the event that an increasing path is present in  $G$  at a time in  $B$ . Note that

$$\mathbb{P}(A_{[0,1]}) \leq \sum_{u \in P} \mathbb{P}(A_{u,[0,1]}) = n^h \mathbb{P}(A_{1,[0,1]}) \quad (2.3)$$

and for any  $\delta \in (0, 1)$  we have

$$A_{1,[0,1]} \subseteq \bigcup_{i=0}^{\lfloor \frac{1}{\delta} \rfloor} A_{1,[i\delta, (i+1)\delta]}$$

Hence by translation invariance,

$$\mathbb{P}(A_{1,[0,1]}) \leq \sum_{i=0}^{\lfloor \frac{1}{\delta} \rfloor} \mathbb{P}(A_{1,[i\delta, (i+1)\delta]}) \leq \frac{1}{\delta} \mathbb{P}(A_{1,[0,\delta]}) \quad (2.4)$$

The aim is now to estimate  $\mathbb{P}(A_{1,[0,\delta]})$ , the probability that path 1 is increasing at some time in  $[0, \delta)$ . For a subset  $S$  of the vertices in the path 1, let  $L_S$  be the event that the subset of vertices specified by  $S$  do not switch during  $[0, \delta)$ . We then write the event  $A_{1,[0,\delta]}$  as

$$A_{1,[0,\delta]} = \bigcup_{k=0}^h \bigcup_{|S|=k} A_{1,[0,\delta]} \cap L_S.$$

Note that the sub-path consisting of those vertices that do not switch must be increasing. Since there are  $\binom{h}{k}$  subsets of size  $k$  and  $\mathbb{P}(A_{1,[0,\delta]} \cap L_S) \leq (1/k!)(1 - e^{-\delta})^{h-k} e^{-\delta k}$  we have that

$$\begin{aligned} \mathbb{P}(A_{1,[0,\delta]}) &\leq \frac{1}{h!} + (1 - e^{-\delta})^h + \sum_{k=1}^{h-1} (1 - e^{-\delta})^{h-k} \frac{1}{k!} \binom{h}{k} \\ &\leq \frac{1}{h!} + \delta^h + \sum_{k=1}^{h-1} \delta^{h-k} \frac{1}{k!} \binom{h}{k} \end{aligned} \quad (2.5)$$

Assuming  $h \geq 3$  and  $\alpha < 1/e$ , set  $\delta = c/h$  where  $c$  is chosen such that  $(1+c)\alpha e < 1$ . We split the above sum according to whether  $1 \leq k \leq \lfloor h/2 \rfloor$  or  $\lfloor h/2 \rfloor \leq k \leq h-1$ . First consider the case  $1 \leq k \leq \lfloor h/2 \rfloor$ . Here we use that Stirling's approximation implies  $k! > (k/e)^k$  and  $\binom{h}{k} \leq (\frac{he}{k})^k$ . Hence,

$$\begin{aligned} \sum_{k=1}^{\frac{h}{2}} \delta^{h-k} \frac{1}{k!} \binom{h}{k} &< \sum_{k=1}^{\lfloor \frac{h}{2} \rfloor} \left(\frac{c}{h}\right)^{h-k} \left(\frac{e}{k}\right)^k \left(\frac{he}{k}\right)^k \\ &\leq \left(\frac{c}{h}\right)^h \sum_{k=1}^{\lfloor \frac{h}{2} \rfloor} \left(\frac{h}{k}\right)^{2k} \left(\frac{e^2}{c}\right)^k \\ &= \left(\frac{c}{h}\right)^h \sum_{k=1}^{\lfloor \frac{h}{2} \rfloor} e^{2k(\log(\frac{he}{\sqrt{c}}) - \log k)} \end{aligned}$$

Note that given  $c \leq 4$  the function  $g(k) = 2k(\log(\frac{he}{\sqrt{c}}) - \log k)$  is increasing on the

interval  $[1, h/2]$  and so the maximum value is  $g(h/2)$ . Therefore

$$\sum_{k=1}^{\frac{h}{2}} \delta^{h-k} \frac{1}{k!} \binom{h}{k} \leq \left(\frac{c}{h}\right)^h \frac{h}{2} \left(\frac{4e^2}{c}\right)^{\frac{h}{2}} = \frac{h}{2} \left(\frac{2ec^{\frac{1}{2}}}{h}\right)^h \quad (2.6)$$

Now consider the case  $\lceil h/2 \rceil \leq k \leq h-1$ . Here we use that

$$\binom{h}{k} = \binom{h}{h-k} \leq \frac{h^{h-k}}{(h-k)!}$$

which gives an upper bound on the required sum of the form

$$\begin{aligned} \sum_{k=\frac{h}{2}}^{h-1} \delta^{h-k} \frac{1}{k!} \binom{h}{k} &\leq \sum_{k=\lceil \frac{h}{2} \rceil}^{h-1} \delta^{h-k} \frac{1}{k^k e^{-k}} \frac{h^{h-k}}{(h-k)^{h-k}} e^{h-k} \\ &= e^h \sum_{k=\lceil \frac{h}{2} \rceil}^{h-1} \delta^{h-k} \left(\frac{h}{h-k}\right)^{h-k} \frac{1}{k^k} \\ &= e^h \sum_{k=\lceil \frac{h}{2} \rceil}^{h-1} e^{(h-k) \log\left(\frac{c}{h-k}\right) - k \log k} \end{aligned}$$

Note that the function  $g(k) = (h-k) \log\left(\frac{c}{h-k}\right) - k \log k$  is maximised at  $k = h/(1+c)$  in the range  $[h/2, h-1]$  provided  $h/2 \leq h/(1+c)$ . For this condition to hold, we choose  $c \leq 1$ . Putting this into the summation above gives

$$e^h \frac{h}{2} e^{-h \log\left(\frac{h}{1+c}\right)} = \frac{h}{2} h^{-h} (e(1+c))^h$$

Therefore combining with (2.6) gives

$$\mathbb{P}(A_{1,(0,\delta)} \cap A_{1,\{0\}}^c) < \frac{h}{2} \left(\frac{2ec^{\frac{1}{2}}}{h}\right)^h + \frac{h}{2} \left(\frac{e(1+c)}{h}\right)^h \leq h \left(\frac{e(1+c)}{h}\right)^h \quad (2.7)$$

Now combining (2.7) with (2.5) and (2.4), recalling that we have chosen  $\delta = c/h$ , gives

$$\mathbb{P}(A_{1,[0,1]}) \leq \left(\frac{h}{c}\right) \mathbb{P}(A_{1,[0,\delta]}) < \left(\frac{h}{c}\right) \frac{1}{h!} + \left(\frac{h}{c}\right) h \left(\frac{e(1+c)}{h}\right)^h \quad (2.8)$$

Combining (2.8) with (2.3) now gives

$$\begin{aligned}
\mathbb{P}(A_{[0,1]}) &\leq \left(\frac{h}{c}\right) \frac{n^h}{h!} + n^h \left(\frac{h}{c}\right) h \left(\frac{e(1+c)}{h}\right)^h \\
&= \left(\frac{1}{c}\right) \left( (\alpha e)^h h \left(\frac{1}{e}\right)^h \frac{h^h}{h!} \right) + h^2 (\alpha h)^h h^{-h} (e(1+c))^h \\
&= \left(\frac{1}{c}\right) \left( (\alpha e)^h h \left(\frac{1}{e}\right)^h \frac{h^h}{h!} \right) + h^2 (\alpha e(1+c))^h
\end{aligned}$$

Note that by Stirling's approximation

$$\left(\frac{1}{e}\right)^h \frac{h^h}{h!} < \frac{1}{2\sqrt{h}} \rightarrow 0,$$

as  $h \rightarrow \infty$ . Additionally, if  $\alpha e < 1$  then  $(\alpha e)^h h \rightarrow 0$  and since we chose  $c$  such that  $\alpha e(1+c) < 1$  we get  $\frac{Ch}{2}(\alpha e(1+c))^h \rightarrow 0$  as  $h \rightarrow \infty$ . Thus  $\alpha < 1/e$  implies there exists<sup>3</sup> a  $\beta > 0$  such that

$$\mathbb{P}(A_{[0,1]}) \leq e^{-\beta h}$$

for  $h$  sufficiently large. □

We can now complete the proof of Part (i) of Theorem 1.1.6:

*Proof of Theorem 1.1.6.* From Proposition 2.3.1 there exists a  $\beta > 0$  such that if  $A_{[s,t]}$  denotes the event that an accessible path appears during the time interval  $[s, t]$  then  $\mathbb{P}(A_{[0,1]}) \leq e^{-\beta h}$  for sufficiently large  $h$ . Hence,

$$\mathbb{P}(A_{[0,\kappa_h]}) \leq \sum_{i=1}^{\kappa_h} \mathbb{P}(A_{[i-1,i]}) \leq \kappa_h e^{-\beta h}.$$

So for any  $\kappa_h > 0$  such that  $\kappa_h e^{-\beta h} \rightarrow 0$  we have  $\mathbb{P}(A_{[0,\kappa_h]}) \rightarrow 0$  as  $h \rightarrow \infty$ . □

We now turn to the second parameterisation in which  $h = en - \beta \log n$ . Here we consider  $\beta < 0$  which is a strongly subcritical regime as the critical value for accessibility percolation in the static case was shown by Chen to be  $\beta = 3/2$  (recall Theorem 1.1.3). Nevertheless we can still provide more information than given by the case  $h = (1/\alpha)n$ ,  $1/\alpha > e$  which was proved above; we show below that as  $-\beta$  gets larger, the time interval on which we are guaranteed to not see an accessible path (with high probability as  $n \rightarrow \infty$ ) grows like  $n^{-\beta-2}$ . To get an upper bound on the probability of existence of an accessible path during a time interval, we declare any vertex that updates to be

---

<sup>3</sup>In particular, any  $\beta < -2 \log(\alpha e(1+c))$ .

*open*. This can be thought of as associating a ‘joker’ symbol to such vertices, meaning we don’t need to consider them when determining whether or not a path is increasing.

*Proof of Part (i) of Theorem 1.1.7.* For a given time interval  $J = [s, t] \subseteq [0, \infty)$ , we say a path  $u \in P$  is *open in J* if the subsequence of vertices whose fitness values do not update during  $J$  is increasing. Formally, if  $0 \leq r \leq h$  vertices do not have their fitness values resampled during  $J$  and  $u(J) = (u_1, \dots, u_r)$  denotes the (increasingly ordered) sequence of indices corresponding to the levels of the non-resampled vertices, the path  $u$  is open during  $J$  if and only if  $X_{u_1}(s) < \dots < X_{u_r}(s)$ . Let  $B_{[s,t]}$  denote the event that there exists an open path in the tree during the interval  $[s, t]$  and note that  $A_{[s,t]} \subseteq B_{[s,t]}$ . Suppose first that we choose  $\kappa_n = -\log(1 - h^{-c})$  for some  $c > 2$ , which is chosen to imply that the probability that a single vertex updates during  $[0, \kappa_n]$  is  $p = h^{-c}$ . The probability that a single path is open is then

$$\sum_{k=0}^h \binom{h}{k} p^{h-k} (1-p)^k \frac{1}{k!}.$$

Note that an upper bound is

$$\sum_{k=0}^h \binom{h}{k} p^{h-k} \frac{1}{k!} \leq \sum_{k=0}^h \binom{h}{k} p^{h-k} \left(\frac{e}{k}\right)^k = \sum_{k=0}^h \binom{h}{k} p^{h-k} \left(\frac{h}{k}\right)^k \left(\frac{e}{h}\right)^k.$$

Now choose  $\varepsilon > 0$  such that  $c > 2 + \varepsilon$  and define  $p' = ph^\varepsilon = h^{-(c-\varepsilon)}$ . Then the right hand side of the above approximation becomes

$$\sum_{k=0}^h \binom{h}{k} (p')^{h-k} \frac{1}{h^{\varepsilon(h-k)}} \left(\frac{h}{k}\right)^k \left(\frac{e}{h}\right)^k = \sum_{k=0}^h \binom{h}{k} (p')^{h-k} f(k) \left(\frac{e}{h}\right)^k$$

where  $f(k) = e^{g(k)}$ ,  $g(k) = -\varepsilon(h-k) \log k + k \log \left(\frac{h}{k}\right)$ . By calculation one can show that  $g(k)$  is increasing for all  $k \in [0, (h^{1+\varepsilon})/e)$  so the term  $f(k)$  takes a maximum value of  $f(h) = e$ . Thus, using the Binomial Theorem we obtain

$$\sum_{k=0}^h (p')^{h-k} f(k) \left(\frac{e}{h}\right)^k \leq e \sum_{k=0}^h \binom{h}{k} (p')^{h-k} \left(\frac{e}{h}\right)^k = e \left(p' + \frac{e}{h}\right)^h$$

Now multiplying by the number of paths,  $n^h$ , we have for sufficiently large  $n$ ,

$$\begin{aligned}
\mathbb{P}(B_{[0, \kappa_n]}) &\leq n^h \left( p' + \frac{e}{h} \right)^h = \left( \frac{en}{h} \right)^h \left( 1 + \frac{p'h}{e} \right)^h = \left( 1 + \frac{\beta \log n}{h} \right)^h \left( 1 + \frac{h^{1-c+\varepsilon}}{e} \right)^h \\
&= \exp \left\{ h \log \left( 1 + \frac{\beta \log n}{h} \right) \right\} \exp \left\{ h \log \left( 1 + \frac{h^{1-c+\varepsilon}}{e} \right) \right\} \\
&\leq n^\beta \exp \left\{ \frac{h^{2-c+\varepsilon}}{e} \right\} \leq n^\beta e^{n^{2-c+\varepsilon}} \\
&\leq 2n^\beta.
\end{aligned} \tag{2.9}$$

where the last inequality follows by our choice of  $\varepsilon$ . Suppose now that  $\frac{\kappa_n}{T_n} = -\log(1 - h^{-c}) \geq h^{-c}$  and note that

$$B_{[0, \kappa_n]} \subseteq \bigcup_{i=1}^{T_n} B_{\left[ \frac{(i-1)\kappa_n}{T_n}, \frac{i\kappa_n}{T_n} \right]}$$

Then by the union bound and (2.9) with  $\kappa_n$  replaced by  $\kappa_n/T_n$ , for large enough  $n$  we get

$$\begin{aligned}
\mathbb{P}(B_{[0, \kappa_n]}) &\leq T_n \mathbb{P} \left( B_{\left[ 0, \frac{\kappa_n}{T_n} \right]} \right) \leq h^c \kappa_n \mathbb{P} \left( B_{\left[ 0, \frac{\kappa_n}{T_n} \right]} \right) \\
&\leq 2h^c \kappa_n n^\beta \leq C \kappa_n n^{\beta+c}
\end{aligned}$$

for some constant  $C > 0$ . Thus, for any  $\beta < 0$ , choosing  $\kappa_n = n^{-\beta-2-\delta}$  for any  $\delta > 0$ , we can choose  $c \in (2, 2 + \delta)$  which implies

$$\mathbb{P}(B_{[0, \kappa_n]}) \rightarrow 0$$

as  $n \rightarrow \infty$ . In particular, for  $\beta$  such that  $\beta + 2 < \eta$  we have no accessible paths almost surely for time intervals of length  $n^\eta$ .  $\square$

## 2.4 Supercritical Case: Increasing paths exist at all times in $[0, 1]$ when $\alpha > \frac{1}{e}$

Here we show that if  $\alpha > 1/e$  then as  $h \rightarrow \infty$  almost-surely for all times in  $[0, 1]$  there exists an accessible path. A key step in the proof below uses a result due to Roberts and Zhao [34]. To state the result, for any  $\varepsilon > 0$  define the set

$$D_\varepsilon = \left\{ (x_1, x_2, \dots, x_h) \in [0, 1]^h : x_j \geq \varepsilon + \frac{(j-1)(1-\varepsilon)}{h} \text{ for all } 1 \leq j \leq h \right\}.$$

On the way to proving Theorem 1.1 in [34] (included here as Theorem 1.1.2 in Section 1.1.2) the authors prove the following:

**Lemma 2.4.1** ([34], display (4.1)). *Consider the House of Cards model on an  $n$ -ary tree of height  $h$  where  $n = \alpha h$  and  $\alpha > 1/e$ . Let  $N_\varepsilon$  denote the number of paths to level  $h$  with fitness in  $I \cap D_\varepsilon$ . Then for all  $\xi \in \log(0, \alpha(1 - \varepsilon)e)$  there exists a constant  $C > 0$  such that for large enough  $h$ ,*

$$\mathbb{P}(N_\varepsilon > e^{\xi h}) \geq \frac{C}{h^3}. \quad (2.10)$$

The main steps of the proof involve a proving a lower bound for  $\mathbb{E}(N_\varepsilon)$  and an upper bound on  $\mathbb{E}(N_\varepsilon^2)$ , which is possible due to the restriction of  $N_\varepsilon$  to count paths in  $I \cap D_\varepsilon$ . These bounds are then combined in a second-moment method to arrive at (2.10). Our result will follow directly from the proof of Theorem 1.1 in Section 4 of [34], in particular using Lemma 4.1 of [34] to slightly adjust Lemma 4.2 in the same paper.

*Proof of Part 3 of Theorem 1.1.6.* For a Borel subset  $B \in \mathcal{B}([0, 1])$ , let  $A_B$  be the event that there is at least one increasing path in the tree for *all* times in  $B$ . Then for any  $\delta \in (0, 1)$  we have

$$\bigcap_{i=0}^{\lfloor \frac{1}{\delta} \rfloor} A_{[i\delta, (i+1)\delta]} \subseteq A_{[0,1]}$$

Thus,

$$\begin{aligned} \mathbb{P}(A_{[0,1]}) &\geq \mathbb{P}\left(\bigcap_{i=0}^{\lfloor \frac{1}{\delta} \rfloor} A_{[i\delta, (i+1)\delta]}\right) = 1 - \mathbb{P}\left(\bigcup_{i=0}^{\lfloor \frac{1}{\delta} \rfloor} A_{[i\delta, (i+1)\delta]}^c\right) \\ &\geq 1 - \sum_{i=0}^{\lfloor \frac{1}{\delta} \rfloor} \mathbb{P}(A_{[i\delta, (i+1)\delta]}^c) \geq 1 - \frac{1}{\delta} \mathbb{P}(A_{[0,\delta]}^c) \end{aligned} \quad (2.11)$$

where we used translation invariance for the last equality. Let  $E_0$  be the event that there are at least  $\frac{n\varepsilon}{10}$  *independent* increasing paths at time 0 where  $\varepsilon$  is chosen such that  $\alpha(1 - \varepsilon)e > 1$ . By independent we mean the paths only share a common vertex at the root. Then we have

$$\mathbb{P}(A_{[0,\delta]}^c) = \mathbb{P}(A_{[0,\delta]}^c | E_0) \mathbb{P}(E_0) + \mathbb{P}(A_{[0,\delta]}^c | E_0^c) \mathbb{P}(E_0^c)$$

Note that  $\mathbb{P}(A_{[0,\delta]}^c | E_0) \leq (1 - e^{-h\delta})^{\frac{n\varepsilon}{10}} \leq (\delta h)^{\frac{n\varepsilon}{10}}$  since at least  $\frac{n\varepsilon}{10}$  vertices of independent paths need to be resampled during  $[0, \delta]$  for all increasing paths to become

non-increasing. Next, we claim that when  $\alpha > 1/e$  there is an  $\eta > 0$  such that

$$\mathbb{P}(E_0) \geq \left(1 - \exp\left(-\frac{n\varepsilon}{4 \cdot 10}\right)\right) (1 - \exp(-\eta h))^{\binom{n\varepsilon}{10}} \quad (2.12)$$

where  $\varepsilon$  is the same as previously. To show this, let  $J_j = [(j-1)\varepsilon/5, j\varepsilon/5)$  for  $1 \leq j \leq 5$  and consider the number of independent increasing subpaths to level 5 in the tree such that the fitness of the  $j$ th vertex in the subpath falls in the interval  $J_j$ . The approach is similar to Lemma 7 in [34], however one difference is that we initially require that there are enough vertices at level 1 with fitness in the interval  $J_1 = [0, \varepsilon/5)$ ; this will give the leftmost factor of (2.12). We will then show exactly as in [34] that there are enough paths continuing for 4 levels further such that the probability that none of them have an accessible extension to level  $h$  is sufficiently small. To carry out this procedure, we use the Chernoff bound (see inequality (2.6) in Theorem 2.1 of [23]), which states that if  $X$  is a binomial random variable then

$$\mathbb{P}\left(X \leq \frac{\mathbb{E}(X)}{2}\right) \leq e^{-\frac{1}{8}\mathbb{E}(X)}. \quad (2.13)$$

Fix an arbitrary level 1 vertex. For  $i > 2$ , let  $M_i$  denote the number of subpaths to level  $i$  from this vertex such that the fitness at level  $j$  lies in the interval  $J_j = [(j-1)\varepsilon/5, j\varepsilon/5)$  where  $1 < j < i$  (recall that we are working at time 0). From this definition,  $M_2 \sim \text{Bin}(j, \varepsilon/5)$ , so from (2.13) we have

$$\mathbb{P}\left(M_2 \leq \frac{n\varepsilon}{10}\right) \leq \exp\left(-\frac{n\varepsilon}{4 \cdot 10}\right) \quad (2.14)$$

For  $i \geq 3$ , we need to account for the number of vertices in the level below whose subpaths from the root can contribute to  $M_i$ . Observe that conditionally on  $M_{i-1}$ ,  $M_i$  has a  $\text{Bin}(nM_{i-1}, \varepsilon/5)$  distribution. For the level 3 vertices we now have

$$\mathbb{E}\left(M_3 \mid M_2 > \frac{n\varepsilon}{10}\right) \geq \frac{n\varepsilon}{10} \cdot \frac{n\varepsilon}{5} = \frac{n^2\varepsilon^2}{5 \cdot 10}$$

Hence applying (2.13) again we get

$$\mathbb{P}\left(M_3 \leq \left(\frac{n\varepsilon}{10}\right)^2 \mid M_2 > \frac{n\varepsilon}{10}\right) \leq \exp\left(-\frac{1}{4}\left(\frac{n\varepsilon}{10}\right)^2\right) \quad (2.15)$$

In a similar way one obtains the inequalities

$$\mathbb{P}\left(M_4 \leq \left(\frac{n\varepsilon}{10}\right)^3 \mid M_3 > \left(\frac{n\varepsilon}{10}\right)^2\right) \leq \exp\left(-\frac{1}{4}\left(\frac{n\varepsilon}{10}\right)^3\right) \quad (2.16)$$



and

$$\mathbb{P}\left(M_5 \leq \left(\frac{n\varepsilon}{10}\right)^4 \mid M_4 > \left(\frac{n\varepsilon}{10}\right)^3\right) \leq \exp\left(-\frac{1}{4}\left(\frac{n\varepsilon}{10}\right)^4\right) \quad (2.17)$$

Therefore summing (2.14), (2.15), (2.16) and (2.17) gives the unconditional upper bound

$$\begin{aligned} \mathbb{P}\left(M_5 \leq \left(\frac{n\varepsilon}{10}\right)^4\right) &\leq \exp\left(-\frac{n\varepsilon}{4 \cdot 10}\right) + \exp\left(-\frac{1}{4}\left(\frac{n\varepsilon}{10}\right)^2\right) \\ &\quad + \exp\left(-\frac{1}{4}\left(\frac{n\varepsilon}{10}\right)^3\right) + \exp\left(-\frac{1}{4}\left(\frac{n\varepsilon}{10}\right)^4\right) \\ &\leq 4 \exp\left(-\frac{n\varepsilon^4}{4 \cdot 10^4}\right) \end{aligned}$$

We now apply Lemma 2.4.1 which in particular implies that if  $\hat{N}_\varepsilon$  denotes the number of accessible paths from the root to level  $h$  with fitness values all above  $\varepsilon$ , there is a constant  $C > 0$  such that  $\mathbb{P}(\hat{N}_\varepsilon > 0) \geq Ch^{-3}$ . Thus the probability that a level 5 vertex with fitness below  $\varepsilon$  has an accessible extension to level  $h$  is bounded below by  $C(h-5)^{-3} \geq C'h^{-3}$  for large enough  $h$  (the supercritical regime still holds for the subtree rooted at level 5 since  $n = \alpha'(h-5)$  where  $\alpha' = \alpha h/(h-5) \geq \alpha > 1/e$ ). Now if  $N$  denotes the total number of accessible paths in the tree of height  $h-1$  rooted at the given level 1 vertex, we have

$$\begin{aligned} \mathbb{P}(N \leq e^{\xi h}) &= \mathbb{P}\left(N \leq e^{\xi h} \mid M_5 > \left(\frac{n\varepsilon}{10}\right)^4\right) + \mathbb{P}\left(M_5 \leq \left(\frac{n\varepsilon}{10}\right)^4\right) \\ &\leq (1 - C'h^{-3})\left(\frac{n\varepsilon}{10}\right)^4 + 4 \exp\left(-\frac{n\varepsilon^4}{4 \cdot 10^4}\right) \\ &\leq \exp\left(-C'h^{-3}\left(\frac{n\varepsilon}{10}\right)^4\right) + 4 \exp\left(-\frac{n\varepsilon^4}{4 \cdot 10^4}\right) \\ &\leq e^{-\eta h} \end{aligned}$$

where  $\eta$  is a constant greater than  $(C'/e)(\alpha\varepsilon/10)^4$  due to the fact that  $n = \alpha h$ . Therefore for any level 1 vertex, the probability that it is connected to an accessible path to level  $h$  is bounded below by  $1 - e^{-\eta h}$ . Now let  $M_1$  denote the number of level 1 vertices with fitness in the interval  $(0, \varepsilon/5)$ . Since  $M_1 \sim \text{Binom}(n, \varepsilon/5)$  we get as in (2.14) that

$$\mathbb{P}\left(M_1 \leq \frac{n\varepsilon}{10}\right) \leq \exp\left(-\frac{n\varepsilon}{4 \cdot 10}\right)$$

Therefore conditioning on the event that  $M_1 > \frac{n\varepsilon}{10}$  we get the lower bound on  $\mathbb{P}(E_0)$  in (2.12). For sufficiently large  $h$  we have

$$(1 - \exp(-\eta h))^{\binom{n\varepsilon}{10}} \geq \left(1 - \left(\frac{10}{n\varepsilon}\right)^3\right)^{\binom{n\varepsilon}{10}}$$

as  $h \rightarrow \infty$ . Now let  $\delta = 1/2h$ . Then we have for large enough  $h$  that

$$\begin{aligned} \mathbb{P}(E_0^c) &\leq 1 - \left(1 - \left(\frac{10}{n\varepsilon}\right)^3\right)^{\left(\frac{n\varepsilon}{10}\right)} + \exp\left(-\frac{n\varepsilon}{4 \cdot 10}\right) \\ &= \left(1 - \left(1 - \left(\frac{10}{n\varepsilon}\right)^3\right)\right) \left(1 + \left(1 - \left(\frac{10}{n\varepsilon}\right)^3\right)^2 + \dots + \left(1 - \left(\frac{10}{n\varepsilon}\right)^3\right)^{\left(\frac{n\varepsilon}{10}\right)-1}\right) \\ &\quad + \exp\left(-\frac{n\varepsilon}{4 \cdot 10}\right) \\ &\leq \left(\frac{10}{n\varepsilon}\right)^2 + \exp\left(-\frac{n\varepsilon}{4 \cdot 10}\right) \end{aligned}$$

It follows that  $\lim_{h \rightarrow \infty} \frac{1}{\delta} \mathbb{P}(E_0^c) = 0$  and hence  $\lim_{h \rightarrow \infty} \mathbb{P}(A_{[0,\delta]}^c) \leq \lim_{h \rightarrow \infty} (\delta h)^{\left(\frac{n\varepsilon}{10}\right)} = 0$  since  $\delta < 1/h$ . Finally, applying this to (2.11) it can be seen that  $\alpha > 1/e$  implies that as  $h \rightarrow \infty$ ,

$$\mathbb{P}(A_{[0,1]}) \rightarrow 1$$

□

## 2.5 Critical case: Second moment bounds

For  $t \in [0, \infty)$  we define  $N(t)$  to be the number of accessible paths present at time  $t$ , formally:

$$N(t) = \sum_{u \in P} \mathbb{1}_{\{X_u(t) \in I\}}.$$

The aim of this section is to prove the following proposition:

**Proposition 2.5.1.** *Let  $\varepsilon \in [0, 1)$  and let  $h = \lfloor en - (\beta + \gamma) \log n \rfloor$ . Now let  $(\kappa_n)_n$  be such that  $\kappa_n \geq 1$  and  $\frac{n^{\frac{3}{2}} \log n}{\kappa_n} \rightarrow 0$  as  $n \rightarrow \infty$ . Then if  $\beta = \gamma = 0$ ,*

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\text{there exists } t \in [0, \kappa_n] \text{ and } u \in P \text{ such that } X_u(t) \in I\right) = 1$$

*We assume that  $\gamma > e\varepsilon - \beta$ . In particular, if  $\kappa_n \ll n^{3/2}$ , then*

$$Q = \mathbb{P}\left(\sup_{t \in [0, \kappa_n]} N(t) > 0\right) \geq \frac{C\kappa_n}{n^{\frac{3}{2}} \log n}$$

*for some  $C > 0$ .*

The following general lemma taken, from [11], will play an important role in our calculations.

**Lemma 2.5.2** ([11], Lemma 2.1). *Let  $J \in \mathbb{N}$  and  $U_1, \dots, U_J \stackrel{i.i.d.}{\sim} \text{Unif}[0, 1]$ . For any*

$\varepsilon \in [0, 1]$  and  $1 \leq k \leq J$  define

$$\psi(k, J, \varepsilon) = \mathbb{P}\left(U_1 \leq \dots \leq U_k, U_j \geq \varepsilon + (1 - \varepsilon)\frac{j-1}{J} \text{ for all } j = 1, \dots, k\right).$$

Then,

$$\psi(k, J, \varepsilon) = \frac{(1 + 1/J)^k (J + 1 - k)}{k!(J + 1)} (1 - \varepsilon)^k.$$

We now outline the general strategy to proving Proposition 2.5.1. Let  $K = \gamma \log n$  for  $\gamma > 0$  and  $h = en - (\beta + \gamma) \log n$  for some  $\beta > 0$ . For  $1 \leq k \leq h$  and  $\varepsilon_n > 0$ , define the set

$$D_{\varepsilon_n, k} = \left\{ (x_1, x_2, \dots, x_k) \in [0, 1]^k : x_j \geq \varepsilon_n + (1 - \varepsilon_n)\frac{j-1}{k} \right\}$$

Set  $\varepsilon_n = \varepsilon \log n/n$ . Let  $P$  denote the set of paths (from the root to a terminal vertex) in the tree, and label the paths of  $P$  according to the position of the terminal vertex, starting from the left (when the tree is embedded in the plane). Thus the path with the left-most terminal vertex will be given the label 1. The number of increasing paths present at time  $t \in [0, \infty)$  that also fall in the set  $D_{\varepsilon_n, h}$  is given by the random variable

$$N_{\varepsilon_n}(t) = \sum_{u \in P} \mathbb{1}_{\{X_u(t) \in I \cap D_{\varepsilon_n, h}\}}.$$

For  $\kappa_n > 0$ , we are interested in the event  $\left\{ \sup_{t \in [0, \kappa_n]} N_{\varepsilon_n}(t) > 0 \right\}$ . For a path  $u \in P$  and  $\kappa_n > 0$  define the random variable

$$Z_u = \int_0^{\kappa_n} \mathbb{1}_{\{X_u(t) \in I \cap D_{\varepsilon_n, h}\}} dt.$$

Observe that

$$\sum_{u \in P} Z_u = \int_0^{\kappa_n} \sum_{u \in P} \mathbb{1}_{\{X_u(t) \in I \cap D_{\varepsilon_n, h}\}} dt = \int_0^{\kappa_n} N_{\varepsilon_n}(t) dt \leq \int_0^{\kappa_n} \sup_{s \in [0, \kappa_n]} N_{\varepsilon_n}(s) dt = \kappa_n \sup_{t \in [0, \kappa_n]} N_{\varepsilon_n}(t).$$

Therefore,

$$\left\{ \sum_{u \in P} Z_u > 0 \right\} \subseteq \left\{ \sup_{t \in [0, \kappa_n]} N_{\varepsilon_n}(t) > 0 \right\}.$$

Using the above inclusion of events and the Paley-Zygmund inequality gives

$$\mathbb{P}\left(\sup_{t \in [0, \kappa_h]} N_{\varepsilon_n}(t) > 0\right) \geq \mathbb{P}\left(\sum_{u \in P} Z_u > 0\right) \geq \frac{\mathbb{E}\left(\sum_{u \in P} Z_u\right)^2}{\mathbb{E}\left(\left(\sum_{u \in P} Z_u\right)^2\right)}. \quad (2.18)$$

For paths  $u, v \in P$ , let  $c(u, v)$  be the number of vertices common to both paths. Expanding the expectations, the right hand side of (2.18) can be written as

$$\mathbb{P}\left(\sup_{t \in [0, \kappa_n]} N(t) > 0\right) \geq \frac{n^{2h} \mathbb{E}(Z_1)^2}{h \sum_{q=0} \sum_{c(u,v)=q} \mathbb{E}(Z_u Z_v)}. \quad (2.19)$$

We aim to show that there is a value  $\beta^*$  such that if  $\beta^* < \beta < \frac{3}{2}$  then

$$\liminf_{h \rightarrow \infty} \mathbb{P}\left(\sup_{t \in [0, \kappa_n]} N_{\varepsilon_n}(t) > 0\right) \geq \frac{1}{\eta}$$

for some value  $\eta$  which using (2.19) will follow if it can be shown that

$$\limsup_{h \rightarrow \infty} \left( \sum_{q=0}^h \sum_{c(u,v)=q} \frac{\mathbb{E}(Z_u Z_v)}{n^{2h} \mathbb{E}(Z_1)^2} \right) \leq \eta. \quad (2.20)$$

We start by calculating the asymptotics of the first moment.

**Lemma 2.5.3.** *We have that if  $\varepsilon_n = \frac{\varepsilon \log n}{n}$*

$$n^h \mathbb{E}(Z_1) = \kappa_n n^{(\beta + \gamma - \varepsilon e - \frac{3}{2})(1+o(1))}.$$

*Proof.* By Lemma 2.5.2, we have using Stirling's approximation that

$$\begin{aligned} n^h \mathbb{E}(Z_1) &= n^h \kappa_n \psi(h, h, \varepsilon_n) \\ &= \kappa_n n^h \left(1 + \frac{1}{h}\right)^h \frac{1}{(h+1)!} (1 - \varepsilon_n)^h \\ &\sim C \kappa_n n^h \frac{e^h}{h^{3/2} h^h} (1 - \varepsilon_n)^h, \end{aligned}$$

for some  $C > 0$ . Using that  $h = en - (\gamma + \beta) \log n$  therefore yields

$$\begin{aligned} n^h \mathbb{E}(Z_1) &= \kappa_n \left(1 - \frac{\beta + \gamma \log n}{e} \frac{1}{n}\right)^{-h} h^{-3/2} e^{-\varepsilon e \log(n)(1+o(1))} \\ &= \kappa_n e^{(\beta + \gamma - \varepsilon e - \frac{3}{2}) \log n (1+o(1))} \\ &= \kappa_n n^{(\beta + \gamma - \varepsilon e - \frac{3}{2})(1+o(1))}, \end{aligned}$$

giving the required result.  $\square$

### 2.5.1 Cases $c(u, v) = 0$ and $c(u, v) = h$

The aim is to obtain a good upper bound on  $\mathbb{E}(Z_u Z_v)$  to prove Proposition 2.5.1. First note that if  $c(u, v) = 0$  then the paths  $u$  and  $v$  are independent so  $\mathbb{E}(Z_u Z_v) = \mathbb{E}(Z_u) \mathbb{E}(Z_v) = \mathbb{E}(Z_1)^2$ . Then, noting that there are  $(n-1)n^{2h-1}$  ordered pairs of independent paths we get that the part of the sum (2.20) corresponding to  $c(u, v) = 0$  is

$$\sum_{c(u,v)=0} \frac{\mathbb{E}(Z_u Z_v)}{n^{2h} \mathbb{E}(Z_1)^2} = \frac{(n-1)n^{2h-1} \mathbb{E}(Z_1)^2}{n^{2h} \mathbb{E}(Z_1)^2} = \frac{n-1}{n} \leq 1 \quad (2.21)$$

Now consider  $c(u, v) = h$ , which implies  $\mathbb{E}(Z_u Z_v) = \mathbb{E}(Z_1^2)$ .

**Lemma 2.5.4.** *If  $\beta + \gamma > \varepsilon e$  and  $\kappa_n \gg n^{\frac{3}{2}}$ , there exists a constant  $C > 0$  such that*

$$\frac{n^h \mathbb{E}(Z_1^2)}{n^{2h} \mathbb{E}(Z_1)^2} \leq \frac{C}{n}$$

for sufficiently large  $n$ .

*Proof.* We can condition on the event that  $k$  vertices are resampled during  $[0, t]$ ,  $A_k$ , which gives

$$\mathbb{E}(Z_1^2) \leq 2\kappa_n \psi(h, h, \varepsilon_n) \sum_{k=0}^h \int_0^{\kappa_n} \mathbb{P}(X_1(t) \in I \cap D_{\varepsilon_n, h} | X_1(0) \in I \cap D_{\varepsilon_n, h}; A_k) \mathbb{P}(A_k) dt \quad (2.22)$$

Isolating the  $k = 0$  term we see that  $\mathbb{P}(A_0) = e^{-ht}$ . Bounding  $\mathbb{P}(X_1(t) \in I \cap D_{\varepsilon_n, h} | X_1(0) \in I \cap D_{\varepsilon_n, h}; A_0)$  above by 1 we then get the  $k = 0$  term (over  $[0, \kappa_n]$ ) in the above summation as:

$$\int_0^{\kappa_n} e^{-ht} dt = \frac{1}{h} (1 - e^{-\kappa_n h}) \leq \frac{1}{h}. \quad (2.23)$$

Note for  $k = h$ , we have that

$$\mathbb{P}(X_1(t) \in I \cap D_{\varepsilon_n, h} | X_1(0) \in I \cap D_{\varepsilon_n, h}; A_h) \leq \psi(h, h, \varepsilon_n).$$

Therefore, if we estimate  $\mathbb{P}(A_h) \leq 1$ , we obtain

$$\begin{aligned} & \frac{n^h 2\kappa_n \psi(h, h, \varepsilon_n)}{n^{2h} \mathbb{E}(Z_1)^2} \int_0^{\kappa_n} \mathbb{P}(X_1(t) \in I \cap D_{\varepsilon_n, h} | X_1(0) \in I \cap D_{\varepsilon_n, h}; A_h) dt \\ & \leq \frac{n^h 2\kappa_n^2 \psi(h, h, \varepsilon_n)^2}{n^{2h} \mathbb{E}(Z_1)^2} = n^{-h}. \end{aligned} \quad (2.24)$$

For the remaining terms note that

$$\mathbb{P}(A_k) = \binom{h}{k} (1 - e^{-t})^k e^{-(h-k)t}$$

and

$$\mathbb{P}(X_u(t) \in I \cap D_{\varepsilon_n} | X_v(0) \in I \cap D_{\varepsilon_n}; A_k) \leq \frac{1}{k!}. \quad (2.25)$$

We will split the  $k$ th summand of the above sum (2.22) into parts according to whether  $0 \leq t \leq 1$  or  $1 \leq t \leq \kappa_n$ . First consider  $1 \leq t \leq \kappa_n$ . In this case we use the bound  $(1 - e^{-t})^k \leq 1$  to get

$$\begin{aligned} \int_1^{\kappa_n} \binom{h}{k} (1 - e^{-t})^k e^{-(h-k)t} dt & \leq \binom{h}{k} \int_1^{\kappa_n} e^{-(h-k)t} dt \leq \binom{h}{k} \frac{1}{h-k} e^{-(h-k)} \\ & \leq \frac{h^k}{k!} \frac{1}{h-k} e^{-(h-k)} \leq \frac{1}{h-k} e^h e^{-(h-k)} = \frac{1}{h-k} e^k. \end{aligned}$$

Therefore, splitting the sum according to whether  $1 \leq k \leq \lfloor h/2 \rfloor$  or  $\lfloor h/2 \rfloor < k \leq h-1$  and applying Stirling's approximation for  $k \geq h/2$  we get

$$\begin{aligned} \sum_{k=1}^{h-1} \frac{1}{h-k} \frac{1}{k!} e^k & \leq \sum_{k=1}^{\lfloor \frac{h}{2} \rfloor} \frac{1}{h-k} \frac{1}{k!} e^k + \sum_{k=\lceil \frac{h}{2} \rceil}^{h-1} \frac{1}{h-k} \frac{1}{k^k} e^k \\ & \leq \frac{2}{h} \sum_{k=1}^{\lfloor \frac{h}{2} \rfloor} \frac{1}{k!} e^k + \sum_{k=\lceil \frac{h}{2} \rceil}^{h-1} \frac{1}{h-k} \frac{1}{k^k} e^{2k} \\ & \leq \frac{2}{h} \sum_{k=1}^{\lfloor \frac{h}{2} \rfloor} \frac{1}{k!} e^k + \sum_{k=\lceil \frac{h}{2} \rceil}^{h-1} \left( \frac{2e^2}{h} \right)^k \\ & \leq \frac{2C_1}{h} + \left( \frac{2e^2}{h} \right)^{\frac{h}{2}}, \end{aligned}$$

for some constant  $C_1 > 0$ . Now consider  $0 \leq t \leq 1$ . For  $k \leq \log h$ , note that for sufficiently large  $h$ ,  $\log h \leq \frac{h}{2}$  so using that  $(1 - e^{-t})^k \leq t^k$  and  $e^{-t(h-k)} \leq e^{-t\frac{h}{2}}$  we have

$$\begin{aligned} \int_0^1 \frac{h^k}{k!} t^k e^{-t\frac{h}{2}} dt &\leq \frac{2}{h} 2^k \int_0^\infty \frac{u^k}{k!} e^{-u} du \\ &= \frac{2^{k+1}}{h} \frac{\Gamma(k+1)}{k!} \leq \frac{2^{k+1}}{h}. \end{aligned}$$

For  $k \geq \log h$  we just bound the integrand, namely  $\mathbb{P}(A_k)$ , by 1 so since we are integrating over  $[0,1]$ , each summand is bounded above by  $\frac{1}{k!}$  using (2.25). Now using the fact that if  $k \geq 4e$  then  $k^{-k}(2e)^k \leq 2^{-k}$  we get by Stirling's approximation,

$$\begin{aligned} \sum_{k=1}^{\log h} \frac{1}{k!} \frac{2^{k+1}}{h} + \sum_{k=\log h+1}^{h-1} \frac{1}{k!} &\leq \frac{1}{h} \sum_{k=1}^{\log h} \frac{1}{k^k} (2e)^k + C_2 \frac{h e^{\log h}}{(\log h)^{\log h}} \\ &\leq \frac{C_1}{h} + C_2 e^{2 \log h - \log(\log h) \log h} \leq \frac{C_3}{h}. \end{aligned}$$

Combining the above with (2.23) and (2.24) and substituting into (2.22), we find that

$$\mathbb{E}(Z_1^2) \leq \frac{C_4 \kappa_n \psi(h, h, \varepsilon_n)}{h}, \quad (2.26)$$

for some  $C_4 > 0$ . Since  $\mathbb{E}(Z_1)^2 = \kappa_n^2 \psi(h, h, \varepsilon_n)^2$  we have

$$\frac{n^h \mathbb{E}(Z_1^2)}{n^{2h} \mathbb{E}(Z_1)^2} \leq \frac{C_5}{h \mathbb{E}(Z_1) n^h}$$

From Lemma 2.5.3, provided  $\beta + \gamma > e\varepsilon$  and if  $\kappa_n \gg n^{\frac{3}{2}}$  we have  $n^h \mathbb{E}(Z_1) \rightarrow \infty$  as  $n \rightarrow \infty$ . Hence under these conditions,

$$\frac{n^h \mathbb{E}(Z_1^2)}{n^{2h} \mathbb{E}(Z_1)^2} \leq \frac{C_6}{n}$$

for sufficiently large  $n$ . □

### 2.5.2 Correlations over short time intervals: $|t - s| < \delta_n$

Now consider  $1 \leq c(u, v) \leq h - 1$ . We consider correlations between times  $t, s \in [0, \kappa_n]$  such that  $|t - s| < \delta_n$  and  $|t - s| \geq \delta_n$  separately:

$$\begin{aligned}
& \sum_{q=1}^{h-1} \sum_{c(u,v)=q} \int_0^{\kappa_n} \int_0^{\kappa_n} \mathbb{P}(X_u(t) \in I \cap D_{\varepsilon_n, h}; X_v(s) \in I \cap D_{\varepsilon_n, h}) ds dt \\
&= \sum_{q=1}^{h-1} \sum_{c(u,v)=q} \int_0^{\kappa_n} \int_0^{\kappa_n} \mathbb{P}(X_u(t) \in I \cap D_{\varepsilon_n, h}; X_v(s) \in I \cap D_{\varepsilon_n, h}) \mathbb{1}_{\{|t-s| < \delta_n\}} ds dt \\
&+ \sum_{q=1}^{h-1} \sum_{c(u,v)=q} \int_0^{\kappa_n} \int_0^{\kappa_n} \mathbb{P}(X_u(t) \in I \cap D_{\varepsilon_n, h}; X_v(s) \in I \cap D_{\varepsilon_n, h}) \mathbb{1}_{\{|t-s| \geq \delta_n\}} ds dt
\end{aligned} \tag{2.27}$$

We have the following lemma for  $|t - s| < \delta_n$ :

**Lemma 2.5.5.** *For any  $\varepsilon < \beta + \gamma$  there is a  $C > 0$  such that*

$$\begin{aligned}
& \sum_{q=1}^{h-1} \sum_{c(u,v)=q} \int_0^{\kappa_n} \int_0^{\kappa_n} \mathbb{P}(X_u(t) \in I \cap D_{\varepsilon_n, h}; X_v(s) \in I \cap D_{\varepsilon_n, h}) \mathbb{1}_{\{|t-s| < \delta_n\}} ds dt \\
& \leq C \kappa_n^{-1} \delta_n h^{3/2} n^{2h} \mathbb{E}(Z_1)^2.
\end{aligned}$$

*Proof.* Let  $u$  and  $v$  be paths such that  $c(u, v) = q$ . Using stationarity, the relevant double integral can be approximated as:

$$\begin{aligned}
& \int_0^{\kappa_n} \int_0^{\kappa_n} \mathbb{P}(X_u(t) \in I \cap D_{\varepsilon_n, h}; X_v(s) \in I \cap D_{\varepsilon_n, h}) \mathbb{1}_{\{|t-s| < \delta_n\}} ds dt \\
& \leq 2\kappa_n \int_0^{\delta_n} \mathbb{P}(X_u(t) \in I \cap D_{\varepsilon_n, h}; X_v(0) \in I \cap D_{\varepsilon_n, h}) dt \\
& \leq 2\kappa_n \int_0^{\delta_n} \mathbb{P}(X_u(t) \in I \cap D_{\varepsilon_n, h} | X_v(0) \in I \cap D_{\varepsilon_n, h}) \mathbb{P}(X_v(0) \in I \cap D_{\varepsilon_n, h}) dt \tag{2.28} \\
& \leq 2\kappa_n \mathbb{P}(X_v(0) \in I \cap D_{\varepsilon_n, h}) \\
& \quad \times \int_0^{\delta_n} \mathbb{P}\left(U_{q+1} < \dots < U_h; U_j \geq \varepsilon_n + (1 - \varepsilon_n) \frac{j-1}{h}, q+1 \leq j \leq h\right) dt,
\end{aligned}$$

where we used the independence of the fitness values from the point where the paths diverge. Define  $\varepsilon'_n = \varepsilon_n + (1 - \varepsilon_n) \frac{q}{h}$  so that  $\frac{(1 - \varepsilon'_n)}{h - q} = \frac{1 - \varepsilon_n}{h}$ . Then we can write the two probabilities using the function  $\psi$  as defined in Lemma 2.5.2 and write the right hand



side of the last display as

$$2\kappa_n \delta_n \psi(h, h, \varepsilon_n) \psi(h - q, h - q, \varepsilon_n').$$

Note that by Lemma 2.5.2

$$\begin{aligned} & \psi\left(h - q, h - q, \varepsilon_n + (1 - \varepsilon_n)\frac{q}{h}\right) \\ &= \left(1 + \frac{1}{h - q}\right)^{h - q} \frac{1}{(h - q + 1)!} \left[1 - \left(\varepsilon_n + (1 - \varepsilon_n)\frac{q}{h}\right)\right]^{h - q} \\ &\leq \frac{e}{(h - q + 1)!} \frac{(1 - \varepsilon_n)^{h - q} (h - q)^{h - q}}{h^{h - q}}. \end{aligned}$$

There are  $(n - 1)n^{2h - q} \leq n^{2h - q}$  ordered pairs of paths with common vertices up to level  $q$ . Therefore, neglecting terms that do not depend on  $q$  we require to calculate the sum

$$\begin{aligned} & \sum_{q=1}^{h-1} n^{-q} \psi\left(h - q, h - q, \varepsilon_n + (1 - \varepsilon_n)\frac{q}{h}\right) \\ &\leq \sum_{q=1}^{h-1} n^{-q} \frac{e}{(h - q + 1)!} \frac{(1 - \varepsilon_n)^{h - q} (h - q)^{h - q}}{h^{h - q}} \\ &\leq \sum_{q=1}^{h-1} n^{-q} \frac{e}{(h - q + 1)(h - q)^{\frac{1}{2}}} \frac{e^{h - q} (1 - \varepsilon)^{h - q}}{h^{h - q}} \\ &\leq \frac{e^{h+1} (1 - \varepsilon)^h}{h^h} \sum_{q=1}^{h-1} \left(\frac{h}{en(1 - \varepsilon_n)}\right)^q \frac{1}{(h - q)^{\frac{3}{2}}} \end{aligned}$$

where we used in the second inequality that by Stirling's approximation  $\frac{1}{k!} \leq \frac{1}{\sqrt{k}} (e/k)^k$ .

Note that

$$\frac{h}{en(1 - \varepsilon_n)} = \frac{1 - (\beta + \gamma)\frac{\log n}{n}}{1 - \varepsilon\frac{\log n}{n}} = 1 - (\beta + \gamma - \varepsilon)\frac{\log n}{n} (1 + o(1)).$$

Now, if  $\varepsilon < \beta + \gamma$ , then  $h < en(1 - \varepsilon_n)$  for  $n$  sufficiently large, so that we have that there exists a constant  $C' > 0$

$$\sum_{q=1}^{h-1} n^{-q} \psi\left(h - q, h - q, \varepsilon_n + (1 - \varepsilon_n)\frac{q}{h}\right) \leq C' \frac{e^{h+1} (1 - \varepsilon_n)^h}{h^h}.$$

Hence, we obtain that the full expression (including the sum over  $q$ ) is bounded from

above by

$$2C'\kappa_n\delta_n n^{2h}\psi(h, h, \varepsilon_n)\frac{e^{h+1}(1-\varepsilon_n)^h}{h^h} \sim C\kappa_n\delta_n n^{2h}h^{3/2}\psi(h, h, \varepsilon_n)^2,$$

for a suitable constant  $C > 0$ , see the proof of Lemma 2.5.3 for a similar calculation. The result now follows using that  $\mathbb{E}(Z_1) = \kappa_n\psi(h, h, \varepsilon_n)$ .  $\square$

### 2.5.3 Correlations over long time intervals: $|t - s| \geq \delta_n$

For  $|t - s| \geq \delta_n$  the idea is to use a ‘Fourier-type’ decomposition to quantify the correlations. For each path  $u \in P$  define the following Boolean function on  $\Omega^m = [0, 1]^m$  where  $m = n^h$  is the number of vertices in  $G$ .

$$f_u(\omega) = \begin{cases} 1 & \text{if } \omega \text{ is such that } u \text{ is an increasing path,} \\ 0 & \text{otherwise.} \end{cases}$$

For each outcome  $\omega \in \Omega$  and  $\varepsilon > 0$ , let  $\omega_\varepsilon$  be a random variable that gives the outcome obtained by resampling the fitness at each vertex in  $G$  independently with probability  $\varepsilon$ . Then the right-most summand of (2.27) is equal to

$$\begin{aligned} & \int_0^{\kappa_n} \int_0^{\kappa_n} \mathbb{P}(X_u(s) \in I; X_v(t) \in I) \mathbb{1}_{\{|t-s| \geq \delta_n\}} ds dt \\ &= \int_0^{\kappa_n} \int_0^{\kappa_n} \mathbb{E}(f_u(\omega)f_v(\omega_{1-e^{-|t-s|}})) \mathbb{1}_{\{|t-s| \geq \delta_n\}} ds dt. \end{aligned} \tag{2.29}$$

### Efron-Stein Decomposition

The Efron-Stein decomposition is a representation of square-integrable random variables defined on product probability spaces, which for indicator functions (also known in this context as Boolean functions) allows us to quantify the influence of the component subspaces on the eventual outcome. The decomposition takes the form of a sum with each summand a random variable associated to the corresponding subspace; if  $f : \Omega^V \rightarrow \{0, 1\}$  for some product space  $\Omega^V$ , we have

$$f = \sum_{S \subseteq V} f_S$$

for a collection of functions  $\{f_S\}_{S \subseteq V}$ . We detail the construction below; this is largely drawn from [31].

Let  $(\Omega, \Sigma, \mu)$  be a probability space and consider the product space  $(\Omega^V, \Sigma^{\otimes V}, \mu^{\otimes S})$ , and denote by  $\omega = (\omega_j, j \in V)$  an element of  $\Omega^V$ .

For  $J \subseteq V$ , we define

$$\omega_J = (\omega_j : j \in J),$$

and for any  $\omega, \omega' \in \Omega^V$  and we define  $(\omega'_{J^c}, \omega_J) \in \Omega^V$  by setting for  $j \in V$

$$(\omega'_{J^c}, \omega_J)_j = \begin{cases} \omega'_j & j \in J^c, \\ \omega_j & j \in J. \end{cases}$$

Informally we define a function that quantifies the expected contribution to  $f$  coming from all subsets of the coordinates in  $J$ . This is the function

$$f^{\subseteq J}(\omega) = \mathbb{E}'(f(\omega'_{J^c}, \omega_J)),$$

where  $\omega'$  is chosen according to  $\mu^{\otimes V}$  and  $\mathbb{E}'$  is the corresponding expectation. In other words,  $f^{\subseteq J}$  is the conditional expectation of  $f$  given the values  $\omega_J$  of  $\omega$  on the subset  $J \subseteq V$ .

The functions  $f^{\subseteq J}$  for  $J \subseteq V$  can be used to construct further functions  $f_J$  which capture the influence on the outcome of  $f$  coming *only* from the subset  $J \subseteq V$  (as opposed to the set of all subsets of  $J$ ).

This is done by finding functions  $\{f_J\}_{J \subseteq V}$  satisfying the condition

$$f^{\subseteq J} = \sum_{K \subseteq J} f_K \tag{2.30}$$

for all subsets  $J \subseteq V$ . Note that the above sum over subsets of  $J$  includes the empty set. Condition (2.30) on its own allows one to iteratively construct the functions  $\{f_J\}_{J \subseteq V}$ . First we define  $f_\emptyset = f^{\subseteq \emptyset} = \mathbb{E}(f)$ . To ensure that (2.30) holds for all singleton subsets  $J = \{j\}$ ,  $j \in V$ , we require that  $f^{\subseteq J} = f_\emptyset + f_J$ , which uniquely determines  $f_J$  as

$$f_J(\omega) = f^{\subseteq J}(\omega) - f_\emptyset(\omega) = \mathbb{E}'(f(\omega') | \omega'_j = \omega_j) - \mathbb{E}(f).$$

Similarly for subsets  $J = \{i, j\}$  containing two elements, the relation (2.30) implies that  $f^{\subseteq J} = f_{\{i\}} + f_{\{j\}} + f_J + f_\emptyset$ , which uniquely determines  $f_J$  since we have already constructed all functions corresponding to singletons as well as the empty set. In

particular,

$$\begin{aligned} f_J(\omega) &= f^{\subseteq J}(\omega) - f_{\{i\}}(\omega) - f_{\{j\}}(\omega) - f_{\emptyset}(\omega) = f^{\subseteq J}(\omega) - f^{\subseteq \{i\}}(\omega) - f^{\subseteq \{j\}}(\omega) + f_{\emptyset}(\omega) \\ &= \mathbb{E}'(f(\omega') | \omega'_i = \omega_i, \omega'_j = \omega_j) - \mathbb{E}'(f(\omega') | \omega'_i = \omega_i) - \mathbb{E}'(f(\omega') | \omega'_j = \omega_j) + \mathbb{E}(f). \end{aligned}$$

Continuing this way by inclusion-exclusion, we define

$$f_S = \sum_{J \subseteq S} (-1)^{|S|-|J|} f^{\subseteq J},$$

and these satisfy (2.30).

We refer to the collection of functions  $(f_J, J \subseteq V)$  as the Efron-Stein decomposition of  $f$ . This decomposition has some notable properties (for a proof see Lemma 8.35 of [31]):

- (P1) For every  $S \subseteq V$ ,  $f_S$  depends only on the coordinates in  $S$ .
- (P2)  $f_T$  and  $f_U$  are orthogonal for  $T \neq U$ , i.e.  $\mathbb{E}(f_T f_U) = 0$ .
- (P3) For all  $\omega \in \Omega^V$  and  $U \subseteq V$ , if  $W \subseteq V$  is such that  $U \not\subseteq W$  then

$$\mathbb{E}(f_U | \omega_W) = 0.$$

### Using the Efron-Stein decomposition to estimate $\mathbb{E}(Z_u Z_v)$

From the Efron-Stein decomposition we can deduce the following lemma which is analogous to Lemma 2.1 in [33]:

**Lemma 2.5.6.** *Fix a product probability space  $(\Omega^n, \otimes_{i=1}^n \Sigma_i, \otimes_{i=1}^n \mu_i)$  and for all  $\omega \in \Omega^n$  let  $\omega_\varepsilon$  be the random variable obtained by resampling each coordinate of  $\omega$  with probability  $\varepsilon \in (0, 1)$ . Then for all  $f, g \in L^2(\Omega^n, \otimes_{i=1}^n \Sigma_i, \otimes_{i=1}^n \mu_i)$  we have*

$$\mathbb{E}(f(\omega)g(\omega_\varepsilon)) = \sum_{T \subseteq [n]} \mathbb{E}(f_T(\omega)g_T(\omega))(1 - \varepsilon)^{|T|} \quad (2.31)$$

*Proof.* Let  $S \subseteq [n]$  be the random set of coordinates whose values are resampled, which is chosen according to the measure  $\mathbb{Q}(S) = \varepsilon^{|S|}(1 - \varepsilon)^{1-|S|}$ . Hence  $|S| \sim \text{Bin}(n, \varepsilon)$ . Note that  $\omega_\varepsilon$  has the same distribution as  $\omega$ . Applying the Efron-Stein expansion gives:

$$\begin{aligned} \mathbb{E}(f(\omega)g(\omega_\varepsilon)) &= \mathbb{E}\left(\sum_{T \subseteq [n]} f_T(\omega) \sum_{U \subseteq [n]} g_U(\omega_\varepsilon)\right) = \sum_{T \subseteq [n]} \sum_{U \subseteq [n]} \mathbb{E}(f_T(\omega)g_U(\omega_\varepsilon)) \\ &= \sum_{T \subseteq [n]} \sum_{U \subseteq [n]} \mathbb{E}\left(\mathbb{E}(f_T(\omega)g_U(\omega_\varepsilon) | \omega)\right) = \sum_{T \subseteq [n]} \sum_{U \subseteq [n]} \mathbb{E}\left(f_T(\omega)\mathbb{E}(g_U(\omega_\varepsilon) | \omega)\right). \end{aligned}$$

Conditioning on the set  $S$  of resampled coordinates,

$$\begin{aligned}\mathbb{E}(g_U(\omega_\varepsilon)|\omega) &= \mathbb{E}(g_U(\omega_\varepsilon)|\omega, U \not\subseteq [n] \setminus S) \mathbb{P}(U \not\subseteq [n] \setminus S) \\ &\quad + \mathbb{E}(g_U(\omega_\varepsilon)|\omega, U \subseteq [n] \setminus S) \mathbb{P}(U \subseteq [n] \setminus S).\end{aligned}\tag{2.32}$$

If  $S$  is such that  $U \not\subseteq [n] \setminus S$  then by property (P3) of the Efron-Stein functions,  $\mathbb{E}(g_U(\omega_\varepsilon)|\omega) = \mathbb{E}(g_U(\omega_\varepsilon)|(\omega_\varepsilon)_i = \omega_i \forall i \in [n] \setminus S) = 0$ . Therefore the expectation in the first summand of (2.32) is zero. Conversely, if  $S$  is such that  $U \subseteq [n] \setminus S$  then by property (P1),  $\mathbb{E}(g_U(\omega_\varepsilon)|\omega, U \subseteq [n] \setminus S) = g_U(\omega)$ . Also,  $\mathbb{P}(U \subseteq [n] \setminus S) = (1 - \varepsilon)^{|U|}$ . Hence,

$$\mathbb{E}(g_U(\omega_\varepsilon)|\omega) = g_U(\omega)(1 - \varepsilon)^{|U|}.$$

Therefore

$$\mathbb{E}(f(\omega)g(\omega_\varepsilon)) = \sum_{T \subseteq [n]} \sum_{U \subseteq [n]} \mathbb{E}(f_T(\omega)g_U(\omega))(1 - \varepsilon)^{|U|}.$$

Note that by property (P1) and the tower property, we can write

$$\mathbb{E}(f_T(\omega)g_U(\omega)) = \mathbb{E}\left(f_T(\omega)\mathbb{E}(g_U(\omega)|\omega_i \forall i \in T)\right) = \mathbb{E}\left(g_U(\omega)\mathbb{E}(f_T(\omega)|\omega_i \forall i \in U)\right).$$

If  $U \not\subseteq T$  then by property (P3),  $\mathbb{E}(g_U(\omega)|\omega_i \forall i \in T) = 0$ ; similarly if  $T \not\subseteq U$  then  $\mathbb{E}(f_T(\omega)|\omega_i \forall i \in U) = 0$  so the above expectation is nonzero only for  $U = T$ . We therefore get

$$\mathbb{E}(f(\omega)g(\omega_\varepsilon)) = \sum_{T \subseteq [n]} \mathbb{E}(f_T(\omega)g_T(\omega))(1 - \varepsilon)^{|T|},$$

which completes the proof of the lemma.  $\square$

### Randomised Algorithm approach

Let  $\{(\Omega_i, \Sigma_i, \mu_i)\}_{i=1}^n$  be an ordered collection of identical probability spaces and  $f : \Omega^n \rightarrow \{0, 1\}$  a measurable function on the corresponding product space. Formally, a randomised algorithm  $A$  determining  $f$  is a finite sequence of indices  $(a_1, a_2, \dots, a_T) \in [n]^T$  for some  $T \in \mathbb{N}$ , chosen consecutively ( $a_i$  after  $a_{i+1}$ ) with the choice of  $a_i$  possibly depending on the outcomes  $\omega_{a_1}, \omega_{a_2}, \dots, \omega_{a_{i-1}}$ , such that knowledge of the values  $\omega_{a_1}, \omega_{a_2}, \dots, \omega_{a_T}$  determines the value of  $f(\omega)$ . We now describe the type of algorithm we use: let  $(\mathcal{T}, \mathcal{F}, \xi)$  be another probability space encoding the randomness used by the algorithm  $A$  in determining  $f$ , and let  $\mathbb{P}$  be the product measure on  $(\Omega^n \times \mathcal{T}, \mathcal{F} \otimes \bigotimes_{i=1}^n \Sigma_i)$ . The algorithm  $A$  at each step reveals one coordinate of the outcome  $\omega \in \Omega^n$ . The algorithm can use the value of the current coordinate as well as

a realisation  $\tau \in \mathcal{T}$  to make decisions on which coordinate to reveal next. Note that this  $\tau$  is fixed for the entire run of the algorithm (e.g. if each decision of the algorithm relies additionally on a coin toss with probability  $p$  then each of the possible coin tosses must be decided beforehand and cannot rely on the realisation of  $\omega \in \Omega^n$  - in this case  $\mathcal{T} = \{0, 1\}^n$  and  $\xi$  is the binomial measure assigning probability  $p^{|x|}(1-p)^{n-|x|}$  to each  $x \in \mathcal{T}$  and  $|x|$  denotes the number of 1s in  $x$ ). The algorithm must be designed in such a way that the value of  $f$  is determined after a finite number of steps. If  $J$  denotes the random set of coordinates revealed by the algorithm  $A$  before  $f$  is determined, and  $U \subseteq [n]$ , the *revelment* of the algorithm  $A$  on the pair  $(U, f)$  is defined as

$$R_A(U, f) = \max_{i \in [n] \setminus U} \mathbb{P}(i \in J),$$

which is the largest probability that a particular coordinate among those outside  $U$  is revealed during the run of the algorithm. The lemma below can be seen as a generalisation of Theorem 2.3 of [33].

**Lemma 2.5.7.** *Let  $A$  be a randomised algorithm revealing a function  $f : \Omega^n \rightarrow \{0, 1\}$  with Efron-Stein decomposition  $f = \sum_{S \subseteq [n]} f_S$  and let  $U \subseteq [n]$ . If  $R_A(U, f)$  denotes the revelment of  $A$  on  $(U, f)$ , then for any integer  $k$  such that  $0 \leq k \leq n - |U|$*

$$\sum_{\substack{|S|=k \\ S \cap U = \emptyset}} \mathbb{E}(f_S(\omega)^2) \leq R_A(U, f) k \mathbb{E}(f(\omega)^2) \quad (2.33)$$

*Proof.* Define

$$h = \sum_{|S|=k, S \cap U = \emptyset} f_S,$$

which means that the Efron-Stein decompositions of  $h$  and  $f$  coincide on subsets  $S \subseteq [n]$  such that  $|S| = k$  and  $S \cap U = \emptyset$ , but otherwise the contributions to  $h$  from all other subsets is zero. Explicitly,

$$h_S(\omega) = \begin{cases} f_S(\omega) & \text{if } |S| = k \text{ and } S \cap U = \emptyset \\ 0 & \text{if } |S| \neq k \text{ or } S \cap U \neq \emptyset \end{cases}$$

which follows by uniqueness of the Efron-Stein decomposition. With this definition, the left-hand side of (2.33) is rewritten as

$$\sum_{\substack{|S|=k \\ S \cap U = \emptyset}} \mathbb{E}(f_S(\omega)^2) = \sum_{S \subseteq [n]} \mathbb{E}(h_S(\omega)^2) = \mathbb{E}(h(\omega)^2). \quad (2.34)$$

The second equality follows by Parseval's formula, which says that given an inner product space  $(H, \langle \cdot, \cdot \rangle)$  and orthogonal vectors  $x_1, x_2, \dots, x_n \in H$ , the sum  $x = \sum_{i=1}^n x_i$  satisfies

$$\langle x, x \rangle = \sum_{i=1}^n \langle x_i, x_i \rangle.$$

This applies in our case since according to property (P2) the Efron-Stein functions form an orthogonal decomposition with respect to the inner product  $\langle g_1, g_2 \rangle = \mathbb{E}(g_1 g_2)$  for functions  $g_1, g_2 \in L^2(\Omega^n)$ . Using this orthogonality, the right-hand side of (2.34) can be rewritten as

$$\begin{aligned} \mathbb{E}(h(\omega)^2) &= \mathbb{E}\left(\sum_{S \subseteq [n]} h_S(\omega) \sum_{S' \subseteq [n]} h_{S'}(\omega)\right) = \mathbb{E}\left(\sum_{\substack{|S|=k \\ S \cap U = \emptyset}} f_S(\omega) \sum_{\substack{|S'|=k \\ S' \cap U = \emptyset}} f_{S'}(\omega)\right) \\ &= \mathbb{E}\left(\sum_{\substack{|S|=k \\ S \cap U = \emptyset}} f_S(\omega) \sum_{S' \subseteq [n]} f_{S'}(\omega)\right) = \mathbb{E}\left(\sum_{S \subseteq [n]} h_S(\omega) \sum_{S' \subseteq [n]} f_{S'}(\omega)\right) \\ &= \mathbb{E}(h(\omega)f(\omega)) \end{aligned}$$

Let  $J = J(\omega, \tau) \subseteq [n]$  denote the (random) set of coordinates revealed by the algorithm  $A$  during its run, and let  $\mathcal{A} = \mathcal{A}(\omega, \tau)$  denote the  $\sigma$ -algebra generated by  $\omega_J = (\omega_j \mid j \in J)$  and  $J$ . Note that  $f$  is  $\mathcal{A}$ -measurable since  $A$  reveals  $f$ . Therefore using the tower property,

$$\begin{aligned} \mathbb{E}(h(\omega)^2) &= \mathbb{E}(h(\omega)f(\omega)) = \mathbb{E}\left(f(\omega)\mathbb{E}(h(\omega)|\mathcal{A})\right) \\ &\leq \mathbb{E}(f(\omega)^2)^{\frac{1}{2}} \mathbb{E}\left(\mathbb{E}(h(\omega)|\mathcal{A})^2\right)^{\frac{1}{2}} \end{aligned} \tag{2.35}$$

where the last inequality is by the Cauchy-Schwarz inequality. We will now condition on  $J$  and  $\omega_J$ , so can treat  $J$  and  $\omega_J$  as fixed. Then, we define the new random variable on  $\Omega^n$

$$h_J(\omega') = h(\omega_J, \omega'_{J^c}).$$

To emphasize the dependency on  $J$  (with in turns depends on  $\omega$  and  $\tau$ ), we write

$$\mathbb{E}(h(\omega') | \mathcal{A})(\omega) = \mathbb{E}^{\omega, \tau}(h_J(\omega')).$$

By the product structure of the underlying measure, we have that  $\omega' = (\omega'_v, v \in V)$  is again a collection of uniformly distributed random variables on  $[0, 1]$ , in other words  $\mathbb{E}^{\omega, \tau}$  corresponds again to the expectation with respect to a product measure on  $\Omega^n$ .

With this notation, we can write the inequality (2.35) as

$$\mathbb{E}(h(\omega)^2) \leq \mathbb{E}(f(\omega)^2)^{\frac{1}{2}} \mathbb{E}\left(\mathbb{E}^{\omega, \tau}(h_J)^2\right)^{\frac{1}{2}}. \quad (2.36)$$

The next step is to bound  $\mathbb{E}(\mathbb{E}^{\omega, \tau}(h_J)^2)$  from above. Note firstly that one can form the Efron-Stein decomposition of  $h_J$  on  $\Omega^n$  with respect to the product measure  $\mathbb{P}^{\omega, \tau}$ . Let  $h_{J,S}$  denote the Efron-Stein function corresponding to  $S \subseteq V$ , so that

$$h_J(\omega') = \sum_{S \subseteq V} h_{J,S}(\omega').$$

As seen in the construction of the Efron-Stein decomposition, the term corresponding to  $S = \emptyset$  is just the expectation (in this case with respect to the measure  $\mathbb{P}^{\omega, \tau}$ ), which gives that

$$h_{J,\emptyset} = \mathbb{E}^{\omega, \tau}(h_J). \quad (2.37)$$

By orthogonality, we get that

$$\mathbb{E}^{\omega, \tau}(h_J(\omega')^2) = \sum_{S \subseteq V} \mathbb{E}^{\omega, \tau}(h_{J,S}(\omega')^2).$$

Therefore, isolating the  $S = \emptyset$  term and rearranging gives by (2.37)

$$(h_{J,\emptyset})^2 = \mathbb{E}^{\omega, \tau}(h_J(\omega')^2) - \sum_{|S| > 0} \mathbb{E}^{\omega, \tau}(h_{J,S}(\omega')^2). \quad (2.38)$$

We will take expectations of this term over  $\mathbb{E}$  and so start by evaluating the expectation of the first term on the right hand side. Using the definition of  $\mathbb{E}^{\omega, \tau}$ , we get that

$$\mathbb{E}(\mathbb{E}^{\omega, \tau}(h_J(\omega')^2)) = \mathbb{E}(\mathbb{E}^{\omega, \tau}(h(\omega_J, \omega'_{J^c})^2)) = \mathbb{E}(\mathbb{E}(h^2 | \mathcal{A})) = \mathbb{E}(h^2) = \sum_{S \subseteq V} \mathbb{E}(h_S^2),$$

using orthogonality (P2) in the last step. Therefore, we obtain by taking expectations in (2.38)

$$\mathbb{E}^{\omega, \tau}(h_J(\omega')^2) = \sum_{S \subseteq V} \mathbb{E}(h_S^2) - \sum_{|S| > 0} \mathbb{E}(\mathbb{E}^{\omega, \tau}(h_{J,S}(\omega')^2)). \quad (2.39)$$

We now claim that

$$h_{J,S}(\omega') = h_S(\omega') \mathbb{1}_{J \cap S = \emptyset}. \quad (2.40)$$

Suppose that  $J \cap S \neq \emptyset$ . First note that since  $h_J$  and therefore also  $h_{J,S}$  only depends on the coordinates in  $J^c$ , we have that  $h_{J,S}(\omega') = \mathbb{E}^{\omega, \tau}(h_{J,S} | \omega'_{J^c})$ . However, since  $S \not\subseteq J^c$ , we can deduce from (P3) that the conditional expectation is zero, so that for



any  $\omega'$

$$h_{J,S}(\omega') = \mathbb{E}^{\omega,\tau}(h_{J,S}|\omega'_{J^c}) = 0. \quad (2.41)$$

Moreover, by the definition of  $h$  and  $h_J$ , we have find that

$$\begin{aligned} h_J(\omega') &= h(\omega_J, \omega'_{J^c}) = \sum_{|S|=k} h_S(\omega_J, \omega'_{J^c}) \\ &= \sum_{|S|=k, S \cap J = \emptyset} h_S(\omega_J, \omega'_{J^c}) + \sum_{|S|=k, S \cap J \neq \emptyset} h_S(\omega_J, \omega'_{J^c}) \\ &= \sum_{|S|=k, S \cap J = \emptyset} h_S(\omega') \end{aligned}$$

where we used in the last step that if  $S \cap J = \emptyset$ , then  $h_S$  does not depend on the coordinates in  $J$  and also (2.41) to see that the second sum is zero. On the other hand, by definition of the Efron-Stein functions, we have that

$$h_J(\omega') = \sum_{S \subseteq V} h_{J,S}$$

By the uniqueness of the Efron-Stein decomposition (see Theorem 8.35 in [31]), we can compare the two decompositions to deduce that the claim (2.40) holds. Combining the decomposition (2.39) with the claim (2.40) and using in the second step that by the definition of  $h_S = 0$  unless  $|S| = k$  and  $S \cap U = \emptyset$  gives

$$\begin{aligned} \mathbb{E}^{\omega,\tau}(h_J(\omega'))^2 &= \sum_{S \subseteq V} \mathbb{E}(h_S^2) - \sum_{|S|>0} \mathbb{E}(\mathbb{E}^{\omega,\tau}(h_S(\omega')^2) \mathbb{1}_{\{J \cap S = \emptyset\}}) \\ &= \sum_{|S|=k, S \cap U = \emptyset} \mathbb{E}(h_S^2) - \sum_{|S|>0} \mathbb{E}(h_S^2) \mathbb{P}(J \cap S = \emptyset) \\ &\leq \sum_{|S|=k, S \cap U = \emptyset} \mathbb{E}(h_S^2) \mathbb{P}(J \cap S \neq \emptyset). \end{aligned}$$

Therefore the inequality (2.36) becomes

$$\mathbb{E}(h(\omega)^2) \leq \mathbb{E}(f(\omega)^2)^{\frac{1}{2}} \left( \sum_{\substack{|S|=k \\ S \cap U = \emptyset}} \mathbb{E}(h_S(\omega)^2) \mathbb{P}(J \cap S \neq \emptyset) \right)^{\frac{1}{2}}. \quad (2.42)$$

The probability that the set of coordinates  $\mathcal{J}$  revealed by  $A$  intersects a given subset  $S$  of size  $k$  with  $S \cap U = \emptyset$  can be bounded uniformly for all such  $S$  in terms of the

revelment  $R_A(U, f)$ :

$$\begin{aligned} \mathbb{P}(\mathcal{J} \cap S \neq \emptyset) &= \mathbb{P}\left(\bigcup_{i \in S} \{i \in \mathcal{J}\}\right) \leq \sum_{i \in S} \mathbb{P}(i \in \mathcal{J}) \leq \sum_{i \in S} \max_{j \in [n] \setminus U} \mathbb{P}(j \in \mathcal{J}) \\ &= \sum_{i \in S} R_A(U, f) = kR_A(U, f). \end{aligned} \quad (2.43)$$

Applying this inequality to (2.42) gives

$$\begin{aligned} \mathbb{E}(h(\omega)^2) &\leq \mathbb{E}(f(\omega)^2)^{\frac{1}{2}} \left( \sum_{\substack{|S|=k \\ S \cap U = \emptyset}} \mathbb{E}(h_S(\omega)^2) kR_A(U, f) \right)^{\frac{1}{2}} \\ &= \mathbb{E}(f(\omega)^2)^{\frac{1}{2}} (kR_A(U, f))^{\frac{1}{2}} \left( \sum_{\substack{|S|=k \\ S \cap U = \emptyset}} \mathbb{E}(h_S(\omega)^2) \right)^{\frac{1}{2}}. \end{aligned} \quad (2.44)$$

Applying Parseval's formula to the sum gives

$$\mathbb{E}(h(\omega)^2) \leq \mathbb{E}(f(\omega)^2)^{\frac{1}{2}} (kR_A(U, f))^{\frac{1}{2}} \mathbb{E}(h(\omega)^2)^{\frac{1}{2}}. \quad (2.45)$$

Dividing by  $\mathbb{E}(h(\omega)^2)^{\frac{1}{2}}$  and squaring both sides gives

$$\mathbb{E}(h(\omega)^2) \leq \mathbb{E}(f(\omega)^2) kR_A(U, f). \quad (2.46)$$

Combining with (2.34) gives the required inequality.  $\square$

### Applying a Randomised algorithm for the modified setup

Here we prove the counterpart to Lemma 2.5.5 for  $|t - s| \geq \delta_h$ .

**Lemma 2.5.8.** *Let  $1 \leq c(u, v) \leq h - 1$  and define the function  $\zeta(\varepsilon, h) = -h \log(1 - \varepsilon)$ .*

*Then there is a  $C > 0$  such that*

$$\begin{aligned} &\sum_{k=1}^{h-1} \sum_{c(u,v)=k} \frac{\int_0^{\kappa_n} \int_0^{\kappa_n} \mathbb{P}(X_u(t) \in I; X_v(s) \in I) \mathbb{1}_{\{|t-s| \geq \delta_n\}} ds dt}{n^{2h} \mathbb{E}(Z_1)^2} \\ &\leq \frac{1}{n} + C\kappa_n^{-1} h^{\frac{3}{2}} \exp\{\zeta(\varepsilon_n, h) - \delta_n\} \end{aligned}$$

*Proof.* For a path  $w \in P$ , let  $\mathcal{U}_w \subseteq [0, 1]^m$  denote the set of possible vectors corresponding to the fitness values of vertices in this path. Note that since  $f_u$  and  $f_v$  only depend on the vertices in their respective paths and both paths agree up to level  $k$ , we can consider both  $f_u$  and  $f_v$  to be Boolean functions on the product space corresponding

to  $\mathcal{U}_u \cup \mathcal{U}_v$  (for  $c(u, v) = k$  we can write  $\mathcal{U}_u \cup \mathcal{U}_v = [0, 1]^V$ , where  $V$  is an ordered set containing the  $2h - k$  vertices comprising the two paths). For subsets  $T \subseteq [m]$  we denote by  $f_u^T$  the function in the Efron-Stein decomposition of  $f_u$  corresponding to the subset  $T$  (and similarly for  $f_v$ ). We specify the ordering of  $V$  by labelling the vertices in path  $u$  by their level (i.e. 1 to  $h$ ) and by attaching label  $h - k + j$  to each vertex in path  $v$  at level  $j$  for  $k + 1 \leq j \leq h$ . Using Lemma 2.5.6 in expression (2.29) from earlier gives:

$$\begin{aligned}
& \int_0^{\kappa_n} \int_0^{\kappa_n} \mathbb{E}(f_u(\omega) f_v(\omega_{1-e^{-|t-s|}})) \mathbb{1}_{\{|t-s| \geq \delta_n\}} ds dt \\
&= \int_0^{\kappa_n} \int_0^{\kappa_n} \sum_{T \subseteq [m]} \mathbb{E}(f_u^T(\omega) f_v^T(\omega)) e^{-|t-s||T|} \mathbb{1}_{\{|t-s| \geq \delta_n\}} ds dt \\
&= \sum_{T \subseteq [m]} \mathbb{E}(f_u^T(\omega) f_v^T(\omega)) \int_0^{\kappa_n} \int_0^{\kappa_n} e^{-|t-s||T|} \mathbb{1}_{\{|t-s| \geq \delta_n\}} ds dt \\
&\leq 2(\kappa_n - \delta_n)^2 \mathbb{E}(f_u^\emptyset(\omega) f_v^\emptyset(\omega)) + \sum_{|T| > 0} \mathbb{E}(f_u^T(\omega) f_v^T(\omega)) 2(\kappa_n - \delta_n) \int_{\delta_n}^{\kappa_n} e^{-t|T|} dt
\end{aligned} \tag{2.47}$$

We can bound the second summand of (2.47) by

$$2\kappa_n \sum_{|T| > 0} \mathbb{E}(f_u^T(\omega) f_v^T(\omega)) \int_{\delta_n}^{\kappa_n} e^{-t|T|} dt. \tag{2.48}$$

Evaluating the integral on the right gives:

$$\int_{\delta_n}^{\kappa_n} e^{-t|T|} dt = \frac{1}{|T|} \left( e^{-\delta_n|T|} - e^{-\kappa_n|T|} \right) \leq \frac{e^{-\delta_n|T|}}{|T|}.$$

Therefore (2.48) is bounded above by

$$2\kappa_n \sum_{|T| > 0} \mathbb{E}(f_u^T(\omega) f_v^T(\omega)) \frac{e^{-\delta_n|T|}}{|T|}. \tag{2.49}$$

Note that if  $T \not\subseteq [k]$ , we claim that  $f_u^T = f_v^T = 0$ . Indeed, let  $T_u \subseteq T$  be the subset of  $T$  containing the vertices of the path  $u$ . Assume  $T \neq \emptyset$  since then  $f_u^T = 0$ . Then since  $T \not\subseteq T_u$  we have by property (P3) that  $\mathbb{E}(f_v^T(\omega) | \omega_{T_u}) = 0$ . Also, since  $f_u^T$  depends only on vertices in  $T_u$ , we have

$$\mathbb{E}(f_u^T(\omega) f_v^T(\omega)) = \mathbb{E}(\mathbb{E}(f_u^T(\omega) f_v^T(\omega) | \omega_{T_u})) = \mathbb{E}(f_u^T(\omega) \mathbb{E}(f_v^T(\omega) | \omega_{T_u})) = 0.$$

By symmetry, the same holds if  $T \not\subseteq T_v$ . If  $T \subseteq T_u \cap T_v = [k]$  then by symmetry we

have  $f_u^T = f_v^T$ , so we can rewrite (2.49) as

$$2\kappa_n \sum_{l=1}^k \frac{e^{-\delta_n l}}{l} \sum_{\substack{|T|=l \\ T \cap [k+1, 2h-k] = \emptyset}} \mathbb{E}(f_u^T(\omega)^2). \quad (2.50)$$

Now we apply a randomised algorithm to determine  $f_u$ . The algorithm will begin at the level  $h$  vertex and continue to check vertices down the path  $u$  until vertex  $k+1$  is reached. If the algorithm has not terminated at this point (this occurs with probability  $\psi(h-k, h-k, \hat{\varepsilon}_n)$  where  $\hat{\varepsilon}_n = \varepsilon_n + (1-\varepsilon_n)(k-1)/h$ ), the algorithm will check the next vertex uniformly from the first  $k$  vertices; we denote this randomly chosen vertex by  $\hat{v}$ . The algorithm will subsequently choose a neighbouring vertex to reveal with probability  $1/2$ . Then the algorithm will continue checking vertices along the direction chosen until either the path stops being increasing or either the root vertex or vertex  $k$  is reached. At this point the algorithm will move back to the uniformly sampled vertex  $\hat{v}$  continue checking in the other direction until either the path becomes non-increasing or the originally chosen vertex is reached. At this point the algorithm will declare that the path is accessible.

The probability of the event  $A_v$  that a vertex  $v \in [k]$  is revealed can be expressed as a sum by conditioning on the distance  $d$  between  $v$  and the  $(h-k+1)$ th chosen vertex  $\hat{v}$  given vertex  $k+1$  is reached. Consider  $\mathbb{P}(A_v | d(v, \hat{v}) = i)$ . We are conditioning on the distance the vertex  $v$  is from the initially chosen vertex,  $\hat{v}$ , and suppose that  $v$  lies at level  $j$  for some  $1 \leq j \leq k$ . Then  $\hat{v}$  lies at level  $j-i$  or  $j+i$ . To reveal vertex  $v$ , the algorithm must accept (i.e. check and not afterwards terminate) all vertices between  $v$  and  $\hat{v}$ , so in particular it must accept at least  $i$  vertices. The probability that  $i$  vertices are accepted is maximised for the vertices at the first  $i$  levels, so for an upper bound we have

$$\mathbb{P}(A_v | d(v, \hat{v}) = i) \leq \psi(i, h, \varepsilon_n).$$

Recalling that we defined  $\hat{\varepsilon}_n = \varepsilon_n + (1 - \varepsilon_n)(k - 1)/h$ , we have

$$\begin{aligned}
\mathbb{P}(A_v) &= \psi(h - k, h - k, \hat{\varepsilon}_n) \sum_{i=0}^{k-1} \mathbb{P}(A_v | d(v, \hat{v}) = i) \mathbb{P}(d(v, \hat{v}) = i) \\
&= \psi(h - k, h - k, \hat{\varepsilon}_n) \sum_{i=0}^{k-1} \mathbb{P}(A_v | d(v, \hat{v}) = i) \mathbb{P}(d(v, \hat{v}) = i) \\
&\leq \psi(h - k, h - k, \hat{\varepsilon}_n) \sum_{i=0}^{k-1} \psi(i, h, \varepsilon_n) \mathbb{P}(d(v, \hat{v}) = i) \\
&\leq \left(1 + \frac{1}{h - k}\right)^{h-k} \frac{(1 - \varepsilon_n)^{h-k}}{(h - k + 1)!} \left(1 - \frac{k - 1}{h}\right)^{h-k} \\
&\quad \times \sum_{i=0}^{k-1} \left(1 + \frac{1}{h}\right)^i \frac{(h + 1 - i)}{i!(h + 1)} (1 - \varepsilon_n)^i \cdot \frac{2}{k - 1} \\
&\leq \frac{e}{2} (h - k + 1)^{-\frac{3}{2}} \left(\frac{(1 - \varepsilon_n)(h - k + 1)}{h(h - k)}\right)^{h-k} \sum_{i=0}^{k-1} \left(\left(1 + \frac{1}{h}\right)(1 - \varepsilon_n)\right)^i \frac{1}{i!} \cdot \frac{2}{k - 1} \\
&\leq \frac{C}{k} (h - k + 1)^{-\frac{3}{2}} \left(\frac{(1 - \varepsilon_n)(h - k + 1)}{h(h - k)}\right)^{h-k} \exp\left\{(1 - \varepsilon_n)\left(1 + \frac{1}{h}\right)\right\}
\end{aligned}$$

where  $C > 0$ . Hence, applying Lemma 2.4.7 with  $U = [k + 1, 2h - k] \cap \mathbb{Z}$ , (2.50) is bounded above by

$$\begin{aligned}
2\kappa_n \sum_{l=1}^k \frac{e^{-\delta_n l}}{l} l \mathbb{E}(f(\omega)^2) R_A(U, f) &\leq C\kappa_h \mathbb{P}(A_v) \mathbb{E}(f(\omega)^2) \sum_{l=1}^k e^{-\delta_n l} \\
&= C\kappa_h \mathbb{P}(A_v) \mathbb{E}(f(\omega)^2) \frac{e^{-\delta_n}}{1 - e^{-\delta_n}} (1 - e^{-\delta_n k})
\end{aligned} \tag{2.51}$$

where we have absorbed all  $O(1)$  terms into the constant  $C > 0$ . We need to multiply the above expression by the number of pairs of paths at level  $k$  (which is bounded above by  $n^{2h-k}$ ), and then sum over the range  $1 \leq k \leq h - 1$ . This gives

$$\begin{aligned}
\frac{C\kappa_h}{h!} \frac{e^{-\delta_n}}{1 - e^{-\delta_n}} n^{2h} \sum_{k=1}^{h-1} \frac{(1 - e^{-\delta_n k}) \mathbb{P}(A_v)}{n^k} &\leq \frac{C\kappa_n e^{-\delta_n}}{h!} n^{2h} \sum_{k=1}^{h-1} \frac{\mathbb{P}(A_v)}{n^k} \\
&\leq \frac{C\kappa_n e^{-\delta_n} n^{2h} e^h}{4h^{h+\frac{1}{2}}} \exp\left\{(1 - \varepsilon_n)\left(1 + \frac{1}{h}\right)\right\} \left(\frac{e(1 - \varepsilon_n)}{h}\right)^h \\
&\quad \times \sum_{k=1}^{h-1} \frac{1}{kn^k} (h - k + 1)^{-\frac{3}{2}} \left(\frac{e(1 - \varepsilon_n)(h - k + 1)}{h(h - k)}\right)^{-k}
\end{aligned} \tag{2.52}$$

Writing out the sum in full we have

$$\begin{aligned}
& \sum_{k=1}^{h-1} \frac{1}{kn^k} (h-k+1)^{-\frac{3}{2}} \left( \frac{e(1-\varepsilon_n)(h-k+1)}{h(h-k)} \right)^{-k} \\
& \leq C \sum_{k=1}^{h-1} \frac{1}{k} \left( \frac{en(1-\varepsilon_n)}{h} \right)^{-k} \frac{1}{(h-k+1)^{3/2}} \\
& \leq C \sum_{k=1}^{h-1} \left( \frac{h}{en(1-\varepsilon_n)} \right)^{h-k} \frac{1}{h-k} \frac{1}{k^{3/2}}
\end{aligned}$$

where we applied Stirling's approximation to obtain the third line. Recall that  $h/en(1-\varepsilon_n) < 1$  provided  $\gamma > e\varepsilon - \beta$  (since  $\varepsilon_n = (\varepsilon \log n)/n$ ). Assuming this condition, the above sum is bounded above by

$$\begin{aligned}
\sum_{k=1}^{h-1} \frac{1}{k^{\frac{3}{2}}} \frac{1}{h-k} & \leq \frac{2}{h} \sum_{k=1}^{h/2} \frac{1}{k^{\frac{3}{2}}} + \frac{1}{h^{\frac{3}{2}}} \sum_{k=h/2}^{h-1} \frac{1}{h-k} \leq \frac{C_1}{h} + \frac{2^{3/2}}{h^{\frac{3}{2}}} \sum_{k=1}^{h/2} \frac{1}{k} \\
& \leq \frac{C_1}{h} + \frac{2^{3/2}}{h^{\frac{3}{2}}} \left( 1 + \log \left( \frac{h}{2} \right) \right) \leq \frac{C_2}{h}
\end{aligned}$$

for large enough  $n$ , where  $C_2 > 2$  is a finite constant. Therefore the combined expression (2.52) is bounded above by

$$\begin{aligned}
& \frac{C\kappa_n e^{-\delta_n} n^{2h} e^h}{4h^{h+\frac{1}{2}}} \cdot \exp \left\{ (1-\varepsilon) \left( 1 + \frac{1}{h} \right) \right\} \left( \frac{e(1-\varepsilon)}{h} \right)^h \cdot \frac{1}{h} \\
& = C_1 \kappa_n e^{-\delta_n} \left( \frac{en}{h} \right)^{2h} \cdot \frac{1}{h^{\frac{3}{2}}} \cdot \exp \left\{ (1-\varepsilon) \left( 1 + \frac{1}{h} \right) + h \log(1-\varepsilon) \right\}
\end{aligned}$$

where the constant  $C$  is implicitly adjusted to account for the constant  $C_2$ . We now need to divide by  $n^{2h} \mathbb{E}(Z_1)^2$ . Recall that

$$n^h \mathbb{E}(Z_1) = C_2 \kappa_n (en h^{-1} (1-\varepsilon_n))^h h^{-\frac{3}{2}}$$

so we obtain the upper bound

$$\begin{aligned}
& \frac{C_1 \kappa_n}{C_2^2 \kappa_n^2} \cdot \frac{h^3}{h^{\frac{3}{2}}} \left( \frac{en(1-\varepsilon_n)}{h} \right)^{-2h} \left( \frac{en}{h} \right)^{2h} \exp \left\{ (1-\varepsilon_n) \left( 1 + \frac{1}{h} \right) + h \log(1-\varepsilon) - \delta_n \right\} \\
& = C_3 \kappa_n^{-1} h^{\frac{3}{2}} \exp \{ \zeta(\varepsilon, h) - \delta_n \}
\end{aligned} \tag{2.53}$$

for a constant  $C_3 > 0$  and  $\zeta(\varepsilon, h) = -h \log(1-\varepsilon)$ . Considering now the terms  $2(\kappa_h -$

$\delta_h) \int_{\delta_h}^{\kappa_h} \mathbb{E}(f_u^\emptyset f_v^\emptyset) \leq \kappa_h^2 \mathbb{P}(X_1(0) \in I \cap D_{\varepsilon_n})^2$  we get (multiplying by  $(n-1)n^{2h-k-1}$  and dividing by  $n^{2h}\mathbb{E}(Z_1)^2$ ) the sum

$$(n-1) \sum_{k=1}^{h-1} \frac{1}{n^{k+1}} \leq \frac{1}{n}.$$

Combining these two bounds gives a total bound of

$$\frac{1}{n} + C_3 \kappa_n^{-1} h^{\frac{3}{2}} \exp\{\zeta(\varepsilon, h) - \delta_n\}$$

□

This last estimate enables us to complete the proof of Proposition 2.5.1:

*Proof of Proposition 2.5.1.* Combining the bound of Lemma 2.5.8 with the expression obtained for the integral up to  $\delta_n$  as well as the terms for  $c(u, v) = 0$  and  $c(u, v) = h$  gives

$$C \kappa_n^{-1} \delta_n h^{3/2} + \frac{C}{n} + 1 + \frac{1}{n} + C \kappa_n^{-1} h^{\frac{3}{2}} \exp\{\zeta(\varepsilon, h) - \delta_n\} \quad (2.54)$$

for some  $C > 0$ . Then choose  $\delta_n$  and  $\kappa_n$  such that  $\delta_n \rightarrow \infty$  and  $\delta_n \kappa_n^{-1} n^{\frac{3}{2}} \rightarrow 0$ . Consider the case in which  $\kappa_n$  is replaced by  $\kappa_n/T_n$  (we still set  $h = en - (\beta + \gamma) \log n$  as in all the above arguments). Then

$$Q = \mathbb{P}\left(\sup_{t \in [0, \kappa_n]} N(t) > 0\right) \geq \frac{1}{1 + C' \max\{\kappa_n^{-1} \delta_n n^{\frac{3}{2}}, \kappa_n^{-1} n^{\frac{3}{2}} e^{\zeta(\varepsilon_n, h) - \delta_n}, n\}}$$

for some  $C' > 0$ . Suppose  $\varepsilon_n > 0$  i.e.  $\varepsilon > 0$ . Note that for any  $\nu > 0$  we can choose  $n$  large enough so that for large  $n$ ,

$$\begin{aligned} \zeta(\varepsilon_n, h) &= -h \log(1 - \varepsilon_n) \leq h \varepsilon_n (1 + \nu) = (en - (\beta + \gamma) \log n) \frac{\varepsilon \log n}{n} (1 + \nu) \\ &= e \varepsilon (1 + \nu) \log n - \frac{\varepsilon (\beta + \gamma) (1 + \nu) (\log n)^2}{n}. \end{aligned}$$

Thus

$$e^{\zeta(\varepsilon_n, h) - \delta_n} \leq D e^{e \varepsilon (1 + \nu) \log n - \delta_n}$$

for some constant  $D > 0$ . Now choose  $\delta_n = e \varepsilon (1 + \nu) \log n$  which causes the second term in the maximum to be smaller than the first by a logarithmic factor and hence

$$Q = \mathbb{P}\left(\sup_{t \in [0, \kappa_n]} N(t) > 0\right) \geq \frac{\kappa_n}{\kappa_n + C'' n^{\frac{3}{2}} \log n} \quad (2.55)$$

for some  $C'' > 0$ . If  $\varepsilon_n = 0$ , i.e.  $\varepsilon = 0$ , then can take  $\delta_n$  to be constant.  $\square$

## 2.6 Extending from the first $K$ levels

The aim of this section is to prove Proposition 2.6.1 below. This works by combining Proposition 2.5.1 proved at the end of the last section with a second moment method which relies on the existence of a large number of accessible paths up to level  $K$  in the tree (using the same  $K$  as before) that do not change for long enough times with high probability.

**Proposition 2.6.1.** *Let  $\beta \in [0, \frac{3}{2})$ . Then for any  $\delta > 0$  and any sequence  $(\kappa_n)_n$  such that  $\kappa_n n^{\beta - \frac{3}{2} + \delta} \rightarrow 0$  as  $n \rightarrow \infty$  we have*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left( \text{there exist } t \in [0, \kappa_n] \text{ and } u \in P \text{ such that } X_u(t) \in I \right) = 1.$$

Now we give an overview of the strategy used to prove Proposition 2.6.1. We split the interval  $[0, \kappa_n]$  into  $T_n \geq 1$  equal sub-intervals; to do this define

$$\tau_i = \left[ \frac{(i-1)\kappa_n}{T_n}, \frac{i\kappa_n}{T_n} \right)$$

for  $1 \leq i \leq T_n$ . The aim is to count paths in the set

$$I_K \cap E_{\varepsilon_n, K} = \left\{ (x_1, x_2, \dots, x_K) \in [0, 1]^K : x_1 < x_2 < \dots < x_K \leq \varepsilon_n; x_j \geq \frac{\varepsilon_n(j-1)}{K} \right\}$$

where  $\varepsilon_n = \varepsilon \log n / n$  for some  $\varepsilon > 0$ . For  $m \in \mathbb{N}$ , a *path to level  $m$*  is a sequence of consecutive vertices beginning at level one and ending at level  $m$ . Given a time-set  $T$ , a path to level  $K$  with fitness values  $X(t) = (x_1(t), x_2(t), \dots, x_K(t)) \in [0, 1]^K$  for  $t \geq 0$  is called  *$\varepsilon$ -good during  $T$*  or just *good during  $T$*  if  $X(t) \in I_K \cap E_{\varepsilon_n, K}$  for all  $t \in T$ . Let  $T(v)$  denote the subtree rooted at vertex  $v$  and for  $x \in [0, 1]$  a sequence of fitness values  $X = (x_1, \dots, x_m) \in [0, 1]^m$  is called  *$x$ -accessible* if  $X \in I_m \cap D_{x, m}$ , i.e. the sequence is increasing and greater than  $x$ . We look at the random variable

$$N = \sum_{|v|=\gamma \log n} \sum_{i=1}^{T_n} \mathbb{1}_{\{v \text{ good during } \tau_i\}} \mathbb{1}_{\{\exists \varepsilon_n\text{-accessible path in } T(v) \text{ during } \tau_i\}}.$$

which counts a subset of the number of accessible paths appearing during the time interval  $\tau_i$ , and apply the second moment method to this (note that  $T(v)$  denotes the subtree rooted at vertex  $v$ ). Note that whenever we write “ $v$  good in  $\tau$ ”, this means that the path from the root to vertex  $v$  is good at the start of time interval  $\tau$  and the



entire path remains un-resampled for the duration of  $\tau$ . In particular, we wish to show that

$$\limsup_{n \rightarrow \infty} \frac{\mathbb{E}(N^2)}{\mathbb{E}(N)^2} \leq 1$$

since by the Paley-Zygmund inequality this will imply

$$\lim_{n \rightarrow \infty} \mathbb{P}(N > 0) = 1. \quad (2.56)$$

Set  $K = \lfloor \gamma \log n \rfloor$  and define the probabilities (equal for any  $v$  with  $|v| = \gamma \log n = K$ )

$$Q = \mathbb{P}(\exists \text{ a } \varepsilon_n\text{-accessible path in } T(v) \text{ during } \tau_i)$$

and

$$Q_{ij} = \mathbb{P}(\exists \text{ a } \varepsilon_n\text{-accessible path in } T(v) \text{ during } \tau_i; \exists \text{ a } \varepsilon_n\text{-accessible path in } T(v) \text{ during } \tau_j).$$

Note firstly that we need to choose the parameters so that  $\mathbb{E}(N) \rightarrow \infty$ . The first moment is

$$\mathbb{E}(N) = n^K T_n Q \mathbb{P}(v \text{ good in } \tau_1)$$

In [11], Chen provides a formula for the probability of being good at a fixed time; we recall this in Chapter 2 as Lemma 2.5.2. It follows from this formula that

$$\mathbb{P}(v \text{ good in } \tau_1) = \psi(K, K, 1 - \varepsilon_n) e^{-\frac{\kappa_n K}{T}}.$$

One can then show that there exist constants  $0 < R_1 < R_2$  such that for large enough  $n$ ,

$$\frac{R_1 n^{\gamma \log(\frac{\varepsilon}{\gamma})}}{(\log n)^{\frac{3}{2}}} \left( \frac{e^{1 - \frac{\kappa_n}{T_n}}}{n} \right)^{\gamma \log n} \leq \mathbb{P}(v \text{ good in } \tau_1) = \frac{R_2 n^{\gamma \log(\frac{\varepsilon}{\gamma})}}{(\log n)^{\frac{3}{2}}} \left( \frac{e^{1 - \frac{\kappa_n}{T_n}}}{n} \right)^{\gamma \log n}, \quad (2.57)$$

and from the proof of Proposition 2.5.1 (with  $\kappa_n$  replaced by  $\kappa_n/T_n$  in (2.55)) we have

$$Q \geq \frac{R_3 \kappa_n}{T_n n^{3/2} \log n}$$

for some constant  $R_3 > 0$ . In particular, we have for large  $n$ ,

$$\begin{aligned}\mathbb{E}(N) &= n^K T_n Q \mathbb{P}(v \text{ good in } \tau_1) \\ &\geq n^{\gamma \log n} T_n \frac{\kappa_n}{T_n n^{3/2} \log n} \frac{R_1 n^{\gamma \log(\frac{\varepsilon}{\gamma})}}{(\log n)^{\frac{3}{2}}} \left( \frac{e^{1 - \frac{\kappa_n}{T_n}}}{n} \right)^{\gamma \log n} \\ &\geq \frac{R_4 \kappa_n n^{\gamma \log(\frac{\varepsilon}{\gamma}) + \gamma}}{n^{3/2} (\log n)^{5/2}}\end{aligned}$$

where  $R_4 = R_3 R_1 > 0$ , and where we used that  $\kappa_n/T_n \rightarrow 0$  as  $n \rightarrow \infty$ . We need the parameters to satisfy

$$\gamma + \beta > e\varepsilon \quad \text{and} \quad \gamma < e\varepsilon$$

where the first inequality is required from the previous sections and the last inequality is needed to ensure that  $\mathbb{E}(N) \rightarrow \infty$ . Note that we require  $\beta > 0$  for these restrictions to hold. We may choose  $\varepsilon$  and  $\gamma$  to satisfy these inequalities as follows: let  $\nu \in (0, 1)$  and choose  $\varepsilon$  such that

$$e\varepsilon = \gamma + (1 - \nu)\beta.$$

To make  $\mathbb{E}(N) \rightarrow \infty$ , for sufficiently large  $\gamma$  we need, as  $n \rightarrow \infty$ ,

$$\kappa_n n^{\gamma \log(\frac{\varepsilon e}{\gamma}) - \frac{3}{2}} = \kappa_n n^{\gamma \log\left(\frac{\gamma + (1 - \nu)\beta}{\gamma}\right) - \frac{3}{2}} \geq \kappa_n n^{(1 - \nu)\beta\left(1 - \frac{C(\beta)}{\gamma}\right) - \frac{3}{2}} \rightarrow \infty$$

where  $C(\beta)$  is a real constant such that  $C(\beta) < 0$  when  $\beta > 0$  and  $C > 0$  when  $\beta < 0$ . Suppose there exists  $\nu > 0$  such that  $\kappa_n n^{\beta - \frac{3}{2} - \nu} \rightarrow \infty$  as  $n \rightarrow \infty$ . We can choose  $\nu \in (0, 1)$  small enough so that  $\beta(1 - \nu)^2 - \frac{3}{2} > 0$ . Now choose  $\gamma$  (and hence also  $\varepsilon$ ) large enough so that

$$n^{(1 - \nu)\beta\left(1 - \frac{C(\beta)}{\gamma}\right) - \frac{3}{2}} < n^{\beta(1 - \nu)^2 - \frac{3}{2}}.$$

It follows that  $\mathbb{E}(N) \rightarrow \infty$  for such a choice of  $\kappa_n$ . To compute the second moment, abbreviate the events  $G(v, i) = \{v \text{ good in } \tau_i\}$  and  $A(v, i) = \{T(v) \text{ accessible in } \tau_i\}$ . Writing  $\mathcal{P}_K$  to denote the set of paths up to level  $K$  and  $J_{v,i} = \mathbb{1}_{G(v,i)} \mathbb{1}_{A(v,i)}$ ,

$$N = \sum_{v \in \mathcal{P}_K} \sum_{i=1}^{T_n} J_{v,i}$$

so using the notation  $I_v = \sum_{i=1}^{T_n} J_{v,i}$  one has

$$\begin{aligned} N^2 &= n^K T_n J_{v,1} + 2n^K \sum_{i=1}^{T_n-1} \sum_{j=i+1}^{T_n} J_{v,i} J_{v,j} + T_n \sum_{q=0}^{K-1} \sum_{c(u,v)=q} J_{v,i} J_{w,i} \\ &+ \sum_{q=0}^{K-1} \sum_{c(u,v)=q} \sum_{i=1}^{T_n-1} \sum_{j=i+1}^{T_n} J_{v,i} J_{w,j} + \sum_{q=0}^{K-1} \sum_{c(u,v)=q} \sum_{i=1}^{T_n-1} \sum_{j=i+1}^{T_n} J_{w,i} J_{v,j}. \end{aligned}$$

Therefore one can write the second moment of  $N$  as

$$\begin{aligned} \mathbb{E}(N^2) &= n^K T_n \mathbb{P}(v \text{ good in } \tau_1) Q \\ &+ 2n^K \sum_{i=1}^{T_n-1} \sum_{j=i+1}^{T_n} \mathbb{P}(v \text{ good in } \tau_i; v \text{ good in } \tau_j) Q_{ij} \\ &+ T_n \sum_{q=0}^{K-1} \sum_{c(v,w)=q} \mathbb{P}(v \text{ good in } \tau_1; w \text{ good in } \tau_1) Q^2 \\ &+ 2 \sum_{q=0}^{K-1} \sum_{c(v,w)=q} \sum_{i=1}^{T_n-1} \sum_{j=i+1}^{T_n} \mathbb{P}(v \text{ good in } \tau_i; w \text{ good in } \tau_j) Q^2. \end{aligned}$$

Note that the first term in the above expansion is just  $\mathbb{E}(N)$ . We now look at each of the remaining terms.

**Lemma 2.6.2** (Second Term). *Set  $p_n = e^{-\frac{\kappa_n}{T_n}}$ . Then there exists a constant  $C > 0$  such that for sufficiently large  $n$ ,*

$$\begin{aligned} &2n^K \sum_{i=1}^{T_n-1} \sum_{j=i+1}^{T_n} \mathbb{P}(v \text{ good in } \tau_i; v \text{ good in } \tau_j) Q_{ij} \\ &\leq \frac{C p_n^K T_n}{\kappa_n \sqrt{n}} \mathbb{E}(N) + \frac{2}{n^K Q} \mathbb{E}(N)^2 + 2 \left( \frac{p_n^K}{1 - p_n^K} \right) \mathbb{E}(N). \end{aligned} \tag{2.58}$$

*Proof.* For the second term we just approximate  $Q_{ij} \leq Q$  giving an upper bound of

$$\begin{aligned} &2n^K Q \sum_{i=1}^{T_n-1} \sum_{j=i+1}^{T_n} \mathbb{P}(v \text{ good in } \tau_i; v \text{ good in } \tau_j) \\ &\leq 2n^K Q \mathbb{P}(v \text{ good in } \tau_1) \sum_{i=1}^{T_n-1} \sum_{j=i+1}^{T_n} \mathbb{P}(v \text{ good in } \tau_j | v \text{ good in } \tau_i) \end{aligned}$$

For  $0 \leq r \leq K$ , let  $R_r$  be the event that  $r$  vertices are resampled during

$$\left[ \frac{i\kappa_n}{T_n}, \frac{(j-1)\kappa_n}{T_n} \right).$$

Note that the events  $\{v \text{ good in } \tau_i\}$  and  $R_r$  are independent for each  $0 \leq r \leq K$ . Hence

$$\begin{aligned} \mathbb{P}(v \text{ good in } \tau_j | v \text{ good in } \tau_i) &= \sum_{r=0}^K \frac{\mathbb{P}(v \text{ good in } \tau_j, v \text{ good in } \tau_i, R_r)}{\mathbb{P}(v \text{ good in } \tau_i)} \\ &= \sum_{r=0}^K \frac{\mathbb{P}(v \text{ good in } \tau_j | v \text{ good in } \tau_i, R_r) \mathbb{P}(v \text{ good in } \tau_i, R_r)}{\mathbb{P}(v \text{ good in } \tau_i)} \\ &= \sum_{r=0}^K \mathbb{P}(v \text{ good in } \tau_j | v \text{ good in } \tau_i, R_r) \mathbb{P}(R_r) \end{aligned}$$

and we aim to get an upper bound on the right-hand side.

*Case 1:*  $1 \leq r \leq K - 1$ .

Note that since  $p_n = e^{-\frac{\kappa_n}{T_n}}$ ,

$$\mathbb{P}(R_r) = \binom{K}{r} (1 - p_n^{j-i-1})^r p_n^{(j-i-1)(K-r)}$$

and that for  $v$  to be good during  $\tau_j$ , we at least need the  $r$  resampled vertices to be such that the resulting sequence of  $K$  fitness values falls in the required set, namely

$$\left\{ (x_1, x_2, \dots, x_K) \in [0, 1]^K : x_1 < x_2 < \dots < x_K \leq \varepsilon_n ; x_j \geq \frac{\varepsilon_n(j-1)}{K} \right\}.$$

This probability is maximised if the lowest  $r$  vertices are resampled, so we use the corresponding probability as an overall upper bound. This is equal to the probability that for all  $1 \leq j \leq r$ , the vertex at distance  $j$  from the root has fitness in the interval  $[(1 - \varepsilon_n) + \frac{\varepsilon_n(j-1)}{K}, 1]$ . Additionally we require the entire path (to level  $K$ ) to remain un-resampled during  $\tau_i$ . We therefore have

$$\mathbb{P}(v \text{ good in } \tau_j | v \text{ good in } \tau_i, R_r) \leq \psi(r, K, 1 - \varepsilon_n) p_n^K,$$

so we aim to get an upper bound on the sum

$$\sum_{r=1}^{K-1} \binom{K}{r} \psi(r, K, 1 - \varepsilon_n) p_n^K (1 - p_n^{j-i-1})^r p_n^{(j-i-1)(K-r)}. \quad (2.59)$$

Note that  $\binom{K}{r} \leq (Ke/r)^r$  and using the result of Chen [11] (see Lemma 2.5.2)

$$\psi(r, K, 1 - \varepsilon_n) \leq \frac{1}{2r^{\frac{1}{2}}} \left(1 + \frac{1}{K}\right)^r \left(1 - \frac{r}{K+1}\right) \left(\frac{e\varepsilon_n}{r}\right)^r. \quad (2.60)$$

Substituting these bounds gives an upper bound on the relevant sum (2.59) of

$$\begin{aligned}
& \frac{p_n^{(j-i)K}}{2} \sum_{r=1}^{K-1} \left(1 + \frac{1}{K}\right)^r \frac{1}{r^{\frac{1}{2}}} \left(1 - \frac{r}{K+1}\right) \varepsilon_n^r \left(\frac{Ke^2}{r^2}\right)^r p_n^{-(j-i-1)r} \\
& \leq \frac{p_n^{(j-i)K}}{2} \sum_{r=1}^{K-1} \left(\frac{(K+1)e^2\varepsilon_n}{r^2}\right)^r p_n^{-(j-i-1)r} \\
& \leq \frac{p_n^{K-1} p_n^{j-i}}{\sqrt{n}} \sum_{r=1}^{K-1} \frac{1}{r^2} < C \frac{p_n^K p_n^{j-i}}{\sqrt{n}}
\end{aligned}$$

where we used that  $\varepsilon_n = \frac{\varepsilon \log n}{n}$  and  $\kappa_n/T_n \rightarrow 0$ , which implies  $(K+1)\frac{e^2}{2}\varepsilon_n < n^{-\frac{1}{2}}$  for sufficiently large  $n$ . Therefore introducing the sums over  $1 \leq i \leq T_n - 1$  and  $i+1 \leq j \leq T_n$  we get

$$C \frac{p_n^K}{\sqrt{n}} \sum_{i=1}^{T_n-1} \sum_{j=i+1}^{T_n} p_n^{(j-i)} = C \frac{p_n^K}{\sqrt{n}} \sum_{i=1}^{T_n-1} \sum_{j=i+1}^{T_n} e^{-(j-i)\frac{\kappa_n}{T_n}} \leq C \frac{p_n^K T_n^2}{\sqrt{n} \kappa_n},$$

where we have adjusted the constant  $C$ . So the eventual upper bound on the second term in the second moment expansion without the  $r = K$  and  $r = 0$  terms is

$$C n^K Q \mathbb{P}(v \text{ good in } \tau_1) p_n^K \frac{T_n^2}{\kappa_n \sqrt{n}} = \frac{C p_n^K T_n}{\kappa_n \sqrt{n}} \mathbb{E}(N).$$

*Case 2:  $r = K$ .*

When  $r = K$  all the vertices in the path are resampled, so in this case  $\mathbb{P}(v \text{ good in } \tau_j | v \text{ good in } \tau_i) = \mathbb{P}(v \text{ good in } \tau_j)$  and hence the contribution to the upper bound of the second term in the second moment expansion is

$$2n^K Q T_n^2 \mathbb{P}(v \text{ good in } \tau_1)^2 = \frac{2}{n^K Q} \mathbb{E}(N)^2.$$

*Case 3:  $r = 0$ .*

An upper bound for the  $r = 0$  term in the sum (2.59) is  $p_n^K p_n^{(j-i-1)K} = p_n^{(j-i)K}$  since in this case no vertices are resampled. Now introducing the sums over time intervals,

we have

$$\sum_{i=1}^{T_n-1} \sum_{j=i+1}^{T_n} p_n^{(j-i)K} \leq \sum_{i=1}^{T_n-1} \sum_{j=1}^{\infty} p_n^{jK} \leq \sum_{i=1}^{T_n-1} p_n^K \frac{1}{1-p_n^K} = T_n p_n^K \frac{1}{1-p_n^K}.$$

Thus the contribution of this case to the upper bound on the second term is bounded above by

$$2n^K Q \mathbb{P}(v \text{ good in } \tau_1) T_n p_n^K \frac{1}{1-p_n^K} = 2\mathbb{E}(N) \left( \frac{p_n^K}{1-p_n^K} \right).$$

Combining the three cases gives the inequality (2.58).  $\square$

**Lemma 2.6.3** (Third Term). *Suppose  $\gamma < e\varepsilon$ . There exists a constant  $C > 0$  such that for sufficiently large  $n$ ,*

$$T_n Q^2 \sum_{q=0}^{K-1} \sum_{c(v,w)=q} \mathbb{P}(v \text{ good during } \tau_1; w \text{ good during } \tau_1) \leq C \frac{(\log n)^{3/2}}{T_n} \mathbb{E}(N)^2 \quad (2.61)$$

*Proof.* The key expression we are considering here is

$$T_n Q^2 \sum_{q=0}^{K-1} \sum_{c(v,w)=q} \mathbb{P}(v \text{ good during } \tau_1; w \text{ good during } \tau_1). \quad (2.62)$$

We already know using results from the estimate (2.57) that there exists a constant  $C_1 > 0$  such that

$$\mathbb{P}(v \text{ good during } \tau_1) \leq \frac{C_1 n^{\gamma \log(\frac{\varepsilon}{\gamma})}}{(\log n)^{\frac{3}{2}}} \left( \frac{e^{1-\frac{\kappa n}{T_n}}}{n} \right)^{\gamma \log n}$$

and we can pull this term outside the double sum in (2.62), which leads to an upper bound of

$$T_n Q^2 \mathbb{P}(v \text{ good during } \tau_1) \sum_{q=0}^{K-1} \sum_{c(v,w)=q} \mathbb{P}(w \text{ good during } \tau_1 | v \text{ good during } \tau_1). \quad (2.63)$$

So to compute the sum we need to compute the conditional probability

$$\mathbb{P}(w \text{ good during } \tau_1 | v \text{ good during } \tau_1)$$

for a given pair of paths to level  $K$  satisfying  $c(u, v) = q$  for any  $q \in [0, K-1] \cap \mathbb{Z}$ . Recall that  $p_n = e^{-\frac{\kappa n}{T_n}}$  is the probability that a single vertex does not update during  $\tau_1$ .

Therefore the relevant conditional probability is bounded above by

$$\mathbb{P}(\text{path to } w \text{ good at start of } \tau_1 \mid \text{path to } v \text{ good at start of } \tau_1) p_n^{K-q}. \quad (2.64)$$

So we now just need to bound the quantity

$$\mathbb{P}(\text{path to } w \text{ good at start of } \tau_1 \mid \text{path to } v \text{ good at start of } \tau_1).$$

Define  $\varepsilon_{n,q} = \varepsilon_n(1 - q/K)$ . To bound the above probability note that if we condition on the first  $q$  vertices having fitness values which make it possible for the whole path to be good, we are interested in the probability that the remaining  $K - q$  vertices are such that the total path is good (for the original  $\varepsilon_n$ ). However since we are dealing with only the upper  $K - q$  vertices whose minimum fitness (to be good) is  $\varepsilon_n(q/K)$ , this is the same as the probability that the  $K - q$  vertices to form a good path with fitness limit  $\varepsilon_{n,q}$  as defined above. Thus the conditional probability in question is bounded above by  $\psi(K - q, K - q, 1 - \varepsilon_{n,q})$ . Writing this in terms of  $\varepsilon_n$  and using Lemma 2.5.2,

$$\begin{aligned} \psi\left(K - q, K - q, 1 - \varepsilon_n\left(1 - \frac{q}{K}\right)\right) &= \left(\frac{K - q + 1}{K - q}\right)^{K-q} \frac{\varepsilon_n^{K-q}}{(K - q + 1)!} \left(\frac{K - q}{K}\right)^{K-q} \\ &\leq \frac{e}{2} \left(\frac{e\varepsilon_n}{K}\right)^K \left(\frac{K}{e\varepsilon_n}\right)^q \frac{1}{(K - q + 1)^{\frac{3}{2}}}. \end{aligned} \quad (2.65)$$

where we used Stirling's approximation in the second line. Now re-introduce the term  $p_n^{K-q}$  from (2.64) and consider the sum over  $q$  (where we also need to include a factor of  $n^{2K-q}$  which is an upper bound for the number of ordered pairs of paths  $(u, v)$  for which  $c(u, v) = q$ ). From the above calculations we have an upper bound on the double sum in (2.63) of

$$C_2 n^{2K} p_n^K \left(\frac{e\varepsilon_n}{K}\right)^K \sum_{q=0}^{K-1} \left(\frac{K}{e\varepsilon_n n p_n}\right)^q \frac{1}{(K - q + 1)^{\frac{3}{2}}} \leq C_3 n^{2K} p_n^K \left(\frac{e\varepsilon_n}{K}\right)^K \sum_{q=0}^{K-1} \left(e^{\frac{\kappa_n}{T_n} - 1} \left(\frac{\gamma}{\varepsilon}\right)\right)^q$$

for constants  $C_2 = e/2$  and  $C_3 = e/(2^{\frac{5}{2}})$ . For the inequality we used that  $K = \gamma \log n$ ,  $\varepsilon_n = \varepsilon \log n/n$  and  $K - q + 1 \geq 2$ . Recall that  $\kappa_n/T_n \rightarrow 0$ , so provided  $\gamma < e\varepsilon$  one can

choose  $n$  large enough so that

$$\sum_{q=0}^{K-1} \left( e^{\frac{\kappa n}{T_n} - 1} \left( \frac{\gamma}{\varepsilon} \right) \right)^q = \frac{1 - \left( e^{\frac{\kappa n}{T_n} - 1} \left( \frac{\gamma}{\varepsilon} \right) \right)^K}{1 - \left( e^{\frac{\kappa n}{T_n} - 1} \left( \frac{\gamma}{\varepsilon} \right) \right)}.$$

Note that

$$\left( e^{\frac{\kappa n}{T_n} - 1} \left( \frac{\gamma}{\varepsilon} \right) \right)^K = n^{\gamma \left( \frac{\kappa n}{T_n} + \log \frac{\varepsilon}{\gamma} \right) - \gamma}$$

and the denominator can be bounded below by a positive constant so we get

$$\sum_{q=0}^{K-1} \left( e^{\frac{\kappa n}{T_n} - 1} \left( \frac{\gamma}{\varepsilon} \right) \right)^q \leq C_4 \left( 1 - n^{\gamma \left( \frac{\kappa n}{T_n} + \log \frac{\varepsilon}{\gamma} \right) - \gamma} \right),$$

for some  $C_4 > 0$ . Now observe that

$$\left( \frac{e\varepsilon n}{K} \right)^K = n^{\gamma \left( 1 + \log \left( \frac{\varepsilon}{\gamma} \right) - \log n \right)}.$$

Thus the combined upper bound on the sum in (2.63) is

$$\begin{aligned} & C_5 n^{2K} p_n^K n^{\gamma \left( 1 + \log \left( \frac{\varepsilon}{\gamma} \right) - \log n \right)} \left( 1 - n^{\gamma \left( \frac{\kappa n}{T_n} + \log \frac{\varepsilon}{\gamma} \right) - K} \right) \\ & \leq C_5 n^{2K} p_n^K n^{\gamma \left( 1 + \log \left( \frac{\varepsilon}{\gamma} \right) - \log n \right)}, \end{aligned}$$

where  $C_5 = C_3 C_4$ . To arrive at an upper bound for the third term in the second moment expansion, we now need to multiply this by

$$T_n Q^2 \mathbb{P}(v \text{ good during } \tau_1) \leq \frac{T_n Q^2 R_2 n^{\gamma \log \left( \frac{\varepsilon}{\gamma} \right)}}{(\log n)^{\frac{3}{2}}} \left( \frac{e^{1 - \frac{\kappa n}{T_n}}}{n} \right)^{\gamma \log n}$$

which gives (noting that  $K = \gamma \log n$  and  $p_n = e^{-\frac{\kappa n}{T_n}}$ ) an upper bound of:

$$C T_n Q^2 (\log n)^{-\frac{3}{2}} n^{2\gamma \left( 1 + \log \left( \frac{\varepsilon}{\gamma} \right) - \frac{\kappa n}{T_n} \right)} \leq C \frac{(\log n)^{3/2}}{T_n} \mathbb{E}(N)^2$$

for some constant  $C > 0$ . □

**Lemma 2.6.4** (Fourth Term). *There exists a constant  $C > 0$  such that for sufficiently*



large  $n$ ,

$$\begin{aligned} & 2Q^2 \sum_{q=0}^{K-1} \sum_{c(v,w)=q} \sum_{i=1}^{T_n-1} \sum_{j=i+1}^{T_n} \mathbb{P}(v \text{ good in } \tau_i; w \text{ good in } \tau_j) \\ & \leq C \frac{(\log n)^{3/2}}{\kappa_n} \mathbb{E}(N)^2 + \frac{C}{n} \mathbb{E}(N)^2 + \mathbb{E}(N)^2. \end{aligned}$$

*Proof.* The fourth term in the expression for  $\mathbb{E}(N^2)$  is

$$\begin{aligned} & 2Q^2 \sum_{q=0}^{K-1} \sum_{c(v,w)=q} \sum_{i=1}^{T_n-1} \sum_{j=i+1}^{T_n} \mathbb{P}(v \text{ good in } \tau_i; w \text{ good in } \tau_j) \\ & = 2Q^2 \mathbb{P}(v \text{ good in } \tau_1) \sum_{q=0}^{K-1} \sum_{c(v,w)=q} \sum_{i=1}^{T_n-1} \sum_{j=i+1}^{T_n} \mathbb{P}(w \text{ good in } \tau_j | v \text{ good in } \tau_i). \end{aligned} \tag{2.66}$$

First consider the outer sum for  $q > 0$ . We will deal with the  $q = 0$  case at the end of this section.

*Case 1:*  $1 \leq q \leq K - 1$ .

For  $1 \leq q \leq K - 1$  and  $0 \leq r \leq q$ , let  $R_r$  be the event that precisely  $r$  of the  $q$  vertices shared between paths  $v$  and  $w$  are resampled between the end of  $\tau_i$  and the start of  $\tau_j$ , i.e. in the interval  $[i\kappa_n/T_n, (j-1)\kappa_n/T_n)$ . Also (as before) let  $p_n = e^{-\frac{\kappa_n}{T_n}}$  denote the avoidance probability for the resampling process of a single vertex in an interval of length  $\kappa_n/T_n$ . Then  $\mathbb{P}(R_r) = \binom{q}{r} (1 - p_n^{j-i-1})^r p_n^{(j-i-1)(q-r)}$  and by independence

$$\mathbb{P}(w \text{ good in } \tau_j | v \text{ good in } \tau_i) = \sum_{r=1}^q \mathbb{P}(w \text{ good in } \tau_j | v \text{ good in } \tau_i \cap R_r) \mathbb{P}(R_r).$$

Expression (2.66) then becomes

$$2Q^2 \sum_{q=0}^{K-1} \sum_{c(v,w)=q} \sum_{r=0}^q \sum_{i=1}^{T_n-1} \sum_{j=i+1}^{T_n} \mathbb{P}(w \text{ good in } \tau_j | v \text{ good in } \tau_i \cap R_r) \mathbb{P}(R_r). \tag{2.67}$$

To bound the probability  $\mathbb{P}(w \text{ good in } \tau_j | v \text{ good in } \tau_i \cap R_r)$ , note that we need the  $r$  resampled vertices to be increasing at the start of time interval  $\tau_j$  and also that the  $K - q$  upper vertices of the path to  $w$  (those vertices not shared with the path to  $v$ ) are such that the whole path to  $w$  is good. The first of these events is bounded above by  $\psi(r, K, 1 - \varepsilon_n)$  (as used in Lemma 2.58 for the second term) and the second is bounded above by  $\psi(K - q, K - q, 1 - \varepsilon_{n,q})$  (as used in Lemma 2.61 for the third

term). Furthermore we need the path to  $w$  to remain un-resampled for the duration of  $\tau_j$ , an event of probability  $p_n^K$ . Therefore we use the bound

$$\mathbb{P}(w \text{ good in } \tau_j | v \text{ good in } \tau_i \cap R_r) \leq \psi(r, K, 1 - \varepsilon_n) \psi(K - q, K - q, 1 - \varepsilon_{n,q}) p_n^K, \quad (2.68)$$

where  $\varepsilon_{n,q} = \varepsilon_n \left(1 - \frac{q}{K}\right)$ . Consider now the case  $1 \leq r \leq q - 1$  (the cases  $r = 0$  and  $r = q$  will be dealt with separately).

*Subcase 1.1:*  $1 \leq r \leq q - 1$ .

Recall from (2.60) that

$$\psi(r, K, 1 - \varepsilon_n) \leq \frac{1}{2r^{\frac{1}{2}}} \left(1 + \frac{1}{K}\right)^r \left(1 - \frac{r}{K+1}\right) \left(\frac{e\varepsilon_n}{r}\right)^r.$$

Also, using the bound  $\binom{q}{r} \leq (qe/r)^r$ , the sum for  $\mathbb{P}(w \text{ good in } \tau_j | v \text{ good in } \tau_i)$  excluding the terms  $r = 0$  and  $r = q$  is bounded above by

$$\begin{aligned} & p_n^K \psi(K - q, K - q, 1 - \varepsilon_{n,q}) \\ & \times \sum_{r=1}^{q-1} \frac{1}{2r^{\frac{1}{2}}} \left(1 + \frac{1}{K}\right)^r \left(1 - \frac{r}{K+1}\right) \left(\frac{e\varepsilon_n}{r}\right)^r \left(\frac{qe}{r}\right)^r (1 - p_n^{j-i-1})^r p_n^{(j-i-1)(q-r)} \\ & \leq C p_n^{K+q(j-i-1)} \psi(K - q, K - q, 1 - \varepsilon_{n,q}) \sum_{r=1}^{q-1} \left(\frac{e^2 q \varepsilon_n}{r^2}\right)^r p_n^{-r(j-i-1)}. \end{aligned} \quad (2.69)$$

Since  $\varepsilon_n = \varepsilon \log n/n$ ,  $K = \gamma \log n$ ,  $0 \leq q \leq K$  and

$$e^2 q \varepsilon_n \leq e^2 K \varepsilon_n = e^2 \varepsilon \gamma (\log n)^2 / n < 1$$

for large enough  $n$ , which gives the bound

$$\sum_{r=1}^{q-1} \left(\frac{e^2 q \varepsilon_n}{r^2}\right)^r p_n^{-r(j-i-1)} \leq C p_n^{-(q-1)(j-i-1)}.$$

This gives an upper bound on the last expression (2.69) of

$$C p_n^{K+q(j-i-1)} \psi(K - q, K - q, 1 - \varepsilon_{n,q}) p_n^{-(q-1)(j-i-1)}.$$

Recall from (2.65) that

$$\psi(K-q, K-q, 1-\varepsilon_{n,q}) \leq \frac{e}{2} \left(\frac{e\varepsilon_n}{K}\right)^K \left(\frac{K}{e\varepsilon_n}\right)^q \frac{1}{(K-q+1)^{\frac{3}{2}}},$$

so bounding  $(K-q+1)$  from below by a constant, we get

$$\mathbb{P}(w \text{ good in } \tau_j | v \text{ good in } \tau_i) \leq C p_n^{K+(j-i-1)} \left(\frac{e\varepsilon_n}{K}\right)^K \left(\frac{K}{e\varepsilon_n}\right)^q.$$

We now need to compute the sum over  $i+1 \leq j \leq T_n$  for each  $1 \leq i \leq T_n-1$ . Only some of the factors in the above expression depend on  $i$  or  $j$  so extracting these, we need to consider the sum

$$\sum_{i=1}^{T_n-1} \sum_{j=i+1}^{T_n} p_n^{(j-i-1)} = e^{\frac{\kappa_n}{T_n}} \sum_{i=1}^{T_n-1} \sum_{j=i+1}^{T_n} e^{-\frac{\kappa_n}{T_n}(j-i)} \leq C \sum_{i=1}^{T_n-1} \frac{1}{1 - e^{-\frac{\kappa_n}{T_n}}} \leq C \frac{T_n^2}{\kappa_n}$$

since  $\kappa_n/T_n \rightarrow 0$  as  $n \rightarrow \infty$ . Next we need to sum over  $q$  which means introducing the counting term  $n^{2K-q}$  (as for the third term). Again extracting terms not depending on  $q$ , we need to compute the following sum for  $1 \leq q \leq K-1$ :

$$\sum_{q=1}^{K-1} \left(\frac{K}{en\varepsilon_n}\right)^q = \sum_{q=1}^{K-1} \left(\frac{\gamma}{e\varepsilon}\right)^q \leq C \left(1 - \left(\frac{\gamma}{e\varepsilon}\right)^K\right) \leq C$$

where we used the assumption  $\gamma < e\varepsilon$ . Hence the upper bound for the terms of (2.67) corresponding to  $1 \leq q \leq K-1$  and  $1 \leq r \leq q-1$  is

$$\begin{aligned} CQ^2 \mathbb{P}(v \text{ good in } \tau_1) n^{2K} p_n^K \left(\frac{e\varepsilon_n}{K}\right)^K \left(\frac{T_n^2}{\kappa_n}\right) &= C \frac{1}{\mathbb{P}(v \text{ good in } \tau_1) \kappa_n} \left(\frac{e\varepsilon_n p_n}{K}\right)^K \mathbb{E}(N)^2 \\ &= C \frac{(\log n)^{3/2}}{n^{\gamma \log(\frac{\varepsilon}{\gamma})} \kappa_n} \left(\frac{en\varepsilon_n e^{-\frac{\kappa_n}{T_n}}}{e^{1-\frac{\kappa_n}{T_n}} K}\right)^K \mathbb{E}(N)^2 \\ &= C \frac{(\log n)^{3/2}}{n^{\gamma \log(\frac{\varepsilon}{\gamma})} \kappa_n} \left(\frac{\varepsilon}{\gamma}\right)^{\gamma \log n} \mathbb{E}(N)^2 \\ &\leq C \frac{(\log n)^{3/2}}{\kappa_n} \mathbb{E}(N)^2 \end{aligned}$$

for some  $C > 0$ .

*Subcase 1.2:  $r = 0$*

The  $r = 0$  case corresponds to no resampling, so the key bound (2.68) reduces to

$$\mathbb{P}(w \text{ good in } \tau_j | v \text{ good in } \tau_i \cap R_0) \leq \psi(K - q, K - q, 1 - \varepsilon_{n,q}) p_n^K.$$

Also  $\mathbb{P}(R_0) = p_n^{(j-i-1)q}$ . So the contribution of the  $r = 0$  term to the full expression for the fourth term (2.66) is bounded above by

$$\begin{aligned} & 4Q^2 p_n^K \mathbb{P}(v \text{ good in } \tau_1) \sum_{q=1}^{K-1} \sum_{c(v,w)=q} \sum_{i=1}^{T_n-1} \sum_{j=i+1}^{T_n} \psi(K - q, K - q, 1 - \varepsilon_{n,q}) p_n^{(j-i-1)q} \\ & \leq \left( \frac{e}{\sqrt{2}} \right) Q^2 p_n^K n^{2K} \left( \frac{e\varepsilon_n}{K} \right)^K \mathbb{P}(v \text{ good in } \tau_1) \sum_{i=1}^{T_n-1} \sum_{j=i+1}^{T_n} \sum_{q=1}^{K-1} \left( \frac{K p_n^{j-i-1}}{en\varepsilon_n} \right)^q \end{aligned}$$

where we introduced the term  $n^{2K-q}$  in the second line and used the estimate

$$\psi(K - q, K - q, 1 - \varepsilon_{n,q}) \leq \frac{e}{2} \left( \frac{e\varepsilon_n}{K} \right)^K \left( \frac{K}{e\varepsilon_n} \right)^q \frac{1}{(K - q + 1)^{\frac{3}{2}}}$$

which we derived earlier in (2.65). Recall that  $K = \gamma \log n$  and  $\varepsilon_n = \varepsilon \log n/n$ , and since we take  $\gamma < e\varepsilon$  we can choose  $n$  large enough so that

$$\sum_{q=1}^{K-1} \left( \left( \frac{\gamma}{\varepsilon} \right) e^{-\frac{\kappa_n}{T_n}(j-i-1)-1} \right)^q \leq e^{-\frac{\kappa_n}{T_n}(j-i)} \sum_{q=1}^{K-1} \left( \left( \frac{\gamma}{e\varepsilon} \right) e^{\frac{\kappa_n}{T_n}-1} \right)^q \leq e^{-\frac{\kappa_n}{T_n}(j-i)} R'$$

Hence, using the values of  $K$ ,  $\varepsilon_n$  and  $p_n$ , the bound on the  $r = 0$  term of (2.66) becomes

$$\begin{aligned} & Q^2 p_n^K n^{2K} \left( \frac{e\varepsilon_n}{K} \right)^K \mathbb{P}(v \text{ good in } \tau_1) \sum_{i=1}^{T_n-1} \sum_{j=i+1}^{T_n} \sum_{q=1}^{K-1} \left( \frac{K p_n^{j-i-1}}{en\varepsilon_n} \right)^q \\ & \leq C n^{2K} \mathbb{P}(v \text{ good in } \tau_1)^2 (\log n)^{3/2} \sum_{i=1}^{T_n-1} \sum_{j=i+1}^{T_n} e^{-\frac{\kappa_n}{T_n}(j-i)}. \end{aligned}$$

As before,

$$\sum_{i=1}^{T_n-1} \sum_{j=i+1}^{T_n} e^{-\frac{\kappa_n}{T_n}(j-i)} = \sum_{i=1}^{T_n-1} \frac{1}{1 - e^{-\frac{\kappa_n}{T_n}}} \leq \frac{T_n^2}{\kappa_n}.$$

Therefore we get the bound

$$C n^{2K} \mathbb{P}(v \text{ good in } \tau_1)^2 (\log n)^{3/2} \frac{T_n^2}{\kappa_n} \leq C \mathbb{E}(N)^2 \frac{(\log n)^{3/2}}{\kappa_n}.$$

*Subcase 1.3:  $r = q$*

In the case  $r = q$ , all the vertices up to the vertex  $q$ , the highest vertex shared between the paths, are resampled. The contribution to the fourth term is bounded above by

$$\begin{aligned}
& \mathbb{P}(w \text{ good in } \tau_1)^2 Q^2 \sum_{q=1}^{K-1} \sum_{c(v,w)=q} \sum_{i=1}^{T_n-1} \sum_{j=i+1}^{T_n} (1 - p_n^{j-i-1})^q \\
& \leq \mathbb{P}(w \text{ good in } \tau_1)^2 n^{2K} Q^2 T_n^2 \sum_{q=1}^{K-1} \frac{1}{n^q} \\
& \leq C \mathbb{P}(w \text{ good in } \tau_1)^2 n^{2K-1} Q^2 T_n^2 = \frac{C}{n} \mathbb{E}(N)^2
\end{aligned}$$

for some  $C > 0$ .

*Case 2:  $q = 0$*

We now look just at the contribution coming from the  $q = 0$  the relevant expression for the fourth term. This becomes

$$\begin{aligned}
& 2Q^2 \mathbb{P}(v \text{ good in } \tau_1) n^K (n^K - 1) \sum_{i=1}^{T_n-1} \sum_{j=i+1}^{T_n} \mathbb{P}(w \text{ good in } \tau_j | v \text{ good in } \tau_i) \\
& = 2Q^2 \mathbb{P}(v \text{ good in } \tau_1) n^K (n^K - 1) \frac{1}{2} T_n (T_n - 1) \mathbb{P}(w \text{ good in } \tau_1) \\
& \leq Q^2 n^{2K} T_n^2 \mathbb{P}(v \text{ good in } \tau_1)^2 = \mathbb{E}(N)^2.
\end{aligned}$$

□

*Proof of Proposition 2.6.1.* Lemmas 2.6.2, 2.6.3 and 2.6.4 give upper bounds for the second, third and fourth terms in the expansion of  $\mathbb{E}(N^2)$  derived in the proof overview. Recall also that the first term is just  $\mathbb{E}(N)$ . Collecting these upper bounds, we get for sufficiently large  $n$ ,

$$\begin{aligned}
\mathbb{E}(N)^2 & \leq \frac{C p_n^K T_n}{\kappa_n \sqrt{n}} \mathbb{E}(N) + \frac{2}{n^K Q} \mathbb{E}(N)^2 + 2 \left( \frac{p_n^K}{1 - p_n^K} \right) \mathbb{E}(N) \\
& \quad + C \frac{(\log n)^{3/2}}{T_n} \mathbb{E}(N)^2 + C \frac{(\log n)^{3/2}}{\kappa_n} \mathbb{E}(N)^2 + \frac{C}{n} \mathbb{E}(N)^2 + \mathbb{E}(N)^2,
\end{aligned}$$

which implies

$$\begin{aligned} \frac{\mathbb{E}(N^2)}{\mathbb{E}(N)^2} &\leq \frac{C p_n^K T_n}{\kappa_n \sqrt{n}} \frac{1}{\mathbb{E}(N)} + \frac{2}{n^K Q} + \left( \frac{p_n^K}{1 - p_n^K} \right) \frac{2}{\mathbb{E}(N)} \\ &\quad + C \frac{(\log n)^{3/2}}{T_n} + C \frac{(\log n)^{3/2}}{\kappa_n} + \frac{C}{n} + 1. \end{aligned}$$

Now suppose  $\kappa_n \gg (\log n)^{3/2}$  and choose  $T_n = (\log n) \kappa_n$ . Note that  $p_n^K \leq 1$  and by our choice of parameters in the proof overview (in particular, choosing  $\gamma$  sufficiently large) we have that  $\mathbb{E}(N) \rightarrow \infty$  as  $n \rightarrow \infty$ . This ensures the first term above goes to zero as  $n \rightarrow \infty$ . The second term also goes to zero since  $Q \geq \frac{\kappa_n}{T_n n^{3/2} \log n}$ . Note also that for any  $\nu > 0$  we have for large enough  $n$ ,

$$\frac{p_n^K}{1 - p_n^K} = \frac{e^{-\gamma \frac{\kappa_n}{T_n} \log n}}{1 - e^{-\gamma \frac{\kappa_n}{T_n} \log n}} \leq \frac{e^{-\gamma \frac{\kappa_n}{T_n} \log n}}{(1 - \nu) \gamma \frac{\kappa_n}{T_n} \log n} \leq \frac{T_n}{(1 - \nu) \gamma \kappa_n \log n}.$$

By our parameter choices in the proof overview we have that  $\mathbb{E}(N) \gg \log n$ , so the third term above goes to zero as  $n \rightarrow \infty$ . Our choices of  $\kappa_n$  and  $T_n$  ensure that the fourth and fifth terms go to zero. Thus we arrive at the result

$$\lim_{n \rightarrow \infty} \frac{\mathbb{E}(N^2)}{\mathbb{E}(N)^2} \leq 1,$$

which by (2.56) gives the statement of the proposition. □

## Chapter 3

# Accessibility Percolation on an Infinite Tree

### 3.1 Introduction

Let  $\mathcal{T} = (V, E)$  be a locally finite, infinite rooted tree where we denote the root vertex by  $o$ . Note that we do not allow the possibility of leaves, i.e. vertices of degree 1. Attach independent random variables uniformly distributed on  $[0,1]$  to each vertex and for each  $h \geq 1$  define the event

$$E_h = \{\exists \text{ an accessible path to level } h\}.$$

We are interested in the existence of an accessible path to infinity, that is, the event

$$E_\infty = \bigcap_{h>0} E_h.$$

Since  $E_{h+1} \subseteq E_h$ ,

$$\mathbb{P}(E_\infty) = \mathbb{P}\left(\bigcap_{h>0} E_h\right) = \lim_{h \rightarrow \infty} \mathbb{P}(E_h).$$

We will use the concept of the *factorial branching number* of a general infinite tree to specify subsets of the set of infinite trees which have accessible paths with positive probability. It should be noted that this quantity does not provide a criterion to completely characterise the accessibility of infinite trees, as one can show that there exist two trees each with factorial branching number equal to 1 such that one has an accessible path with positive probability, while the other has no accessible path almost surely. To define the factorial branching number, recall that a *cutset*  $\pi \subseteq E$  is a set of

edges such that all paths from the root to infinity pass through  $\pi$  and such that the removal of any edge from  $\pi$  produces a path to infinity not passing through  $\pi$ . Denote by  $\Pi = \Pi(\mathcal{T})$  the set of all cutsets of  $\mathcal{T}$  and for an edge  $e \in E$  let  $|e|$  denote the level of  $e$ . The *factorial branching number* of  $\mathcal{T}$  is then defined as the quantity

$$\text{br}_f(\mathcal{T}) = \sup \left\{ \lambda > 0 : \inf_{\pi \in \Pi} \sum_{e \in \pi} (|e|!)^{-\lambda} > 0 \right\}. \quad (3.1)$$

The rest of the chapter is structured as follows: in Section 3.2 we introduce the concepts we will need from the theory of electrical networks on trees as developed by Lyons and Peres among others (see [29] for a detailed exposition). In Section 3.3 we define a form of percolation such that an infinite connected component of open edges implies the existence of an infinite accessible path. Using this setup and the theory developed in Section 3.2, we then prove Propositions 3.3.1 and 3.3.2, which cover the supercritical ( $\text{br}_f(\mathcal{T}) > 1$ ) and subcritical ( $\text{br}_f(\mathcal{T}) < 1$ ) cases respectively. In Section 3.4 we turn to an investigation of the critical case and show that for spherically symmetric trees such that the number of descendant vertices grows linearly with the level of the parent vertex, no accessible paths exist almost surely. In Section 3.5, we look at how the ideas from Section 3.3 as well as the paper of Khoshnevisan [25] could be used to prove a result on exceptional times of accessibility for general trees in the critical case.

## 3.2 Electrical Networks

We now recall various useful concepts from the electrical network approach to probability on trees, as detailed in the book by Lyons and Peres [29]. We use the notation that for a vertex  $x \in V$ ,  $e(x)$  denotes the edge adjacent to  $x$  closest to the root, and for an edge  $e \in E$ , we label the vertices on either end by the pair  $(e^-, e^+)$ , where  $e^-$  is closest to the root. Also for each vertex  $v \in V$  we let  $N(v)$  denote its set of neighbouring vertices, and for all  $w \in N(v)$  we let  $(v, w)$  denote the edge joining  $v$  and  $w$ . Fix a non-negative function  $c : E \rightarrow [0, \infty)$  which assigns weights called *conductances*  $c(e)$  to each edge  $e \in E$ . We now define the *effective conductance* between two (not necessarily adjacent) vertices  $x, y \in V$ . Let  $(W_n)_{n \geq 0}$  be the nearest neighbour random walk started at  $x$  such that for each  $v$ ,

$$\mathbb{P}(W_{n+1} = w \mid W_n = v) = \frac{c(v, w)}{\sum_{w' \in N(v)} c(v, w')}.$$

Let  $\mathbb{P}[x \longrightarrow y]$  denote the probability that the random walk  $(W_n)_{n \geq 0}$  started at  $x$  hits  $y$  before returning to  $x$ . The effective conductance between  $x$  and  $y$ , denoted  $\mathcal{C}(x \longleftrightarrow y)$ ,



is defined as

$$\mathcal{C}(x \longleftrightarrow y) = \mathbb{P}[x \longrightarrow y] \sum_{w \in N(x)} c(x, w) = \mathbb{P}[y \longrightarrow x] \sum_{w \in N(y)} c(y, w),$$

where the second equality follows because  $(W_n)_{n \geq 0}$  is a reversible Markov chain. As we are working on a tree, the effective conductance between vertices  $x$  and  $y$  in  $V$  can be found via the summation rule for conductors in series,

$$\frac{1}{\mathcal{C}(x \longleftrightarrow y)} = \sum_{e \in E_{x,y}} \frac{1}{c(e)},$$

where  $E_{x,y}$  is the set of edges in the unique path joining  $x$  and  $y$ . In addition, the effective conductance between a vertex  $x \in V$  and infinity, denoted  $\mathcal{C}(x \longleftrightarrow \infty)$ , is defined as follows. Let  $(\pi_n)_n$  be a sequence of cutsets of the subtree rooted at  $x$ ,  $\mathcal{T}_x$ , such that for all  $n \in \mathbb{N}$  any path from a vertex in  $\pi_n$  to a vertex in  $\pi_{n+2}$  must pass through  $\pi_{n+1}$ . Note that the cutsets  $(\pi_n)_n$  generate a sequence of finite subtrees  $T_n = (V_n, E_n)$  satisfying  $T_n \subseteq T_{n+1}$  for which the set of terminal vertices of  $T_n$  is equal to  $\{e^+ : e \in \pi_n\}$ . Additionally, the subtrees  $(T_n)_n$  exhaust  $\mathcal{T}_x$ , that is,  $\mathcal{T}_x = \bigcup_{n=1}^{\infty} T_n$ . Then Remark 2.10 in Lyons and Peres [29] shows that

$$\frac{1}{\mathcal{C}(x \longleftrightarrow \infty)} = \sum_{n=1}^{\infty} \left( \frac{1}{\sum_{e \in \pi_n} c(e)} \right).$$

We can also consider functions  $\theta : E \rightarrow [0, \infty)$  such that for any edge  $e$  with  $D_e$  its set of descendant edges at the next level we have

$$\theta(e) = \sum_{d \in D_e} \theta(d).$$

Any such function is called a *flow*. The *strength* of a flow  $\theta$  is defined as the total flowing into the tree from the root, that is,

$$\text{Strength}(\theta) = \sum_{e \in E : e^- = o} \theta(e). \quad (3.2)$$

For a given choice of conductances  $(c(e))_{e \in E}$ , the *energy* of a flow  $\theta$  is defined by

$$\mathcal{E}(\theta) = \sum_{e \in E} \frac{\theta(e)^2}{c(e)}. \quad (3.3)$$

The primary concern later on will be to find, for a particular choice of conductances  $(c(e))_{e \in E}$ , a nonzero flow of finite energy. Another useful fact is Thomson's principle (see Chapter 4, Section 2.4 of [29]), which states that

$$\mathcal{C}(o \longleftrightarrow \infty) = \frac{1}{\inf \{ \mathcal{E}(\theta) : \theta \text{ is a flow of strength } 1 \}} \quad (3.4)$$

Consider the following assignment of numbers in  $(0,1)$  to elements of  $E$ . For each  $n \in \mathbb{N}$ , let  $p_n$  be the probability that an edge  $e$  at distance  $n$  from the root is open. For this assignment, we are interested in  $\mathbb{P}(o \longleftrightarrow \infty)$ , the probability that there exists a set of open edges consisting of a path from the root to infinity. To investigate this probability, the first step is to generate a specific set of conductances  $(c(e))_{e \in E}$  associated to this assignment. Following the approach of Lyons and Peres, this is done by first defining a *gauge*  $\Psi : V \rightarrow [0, \infty]$  on the vertices of  $\mathcal{T}$ , given by

$$\Psi(x) = \frac{1}{\mathbb{P}(o \longleftrightarrow x)} = \prod_{i=1}^{|x|} \frac{1}{p_i} \quad \text{for } x \in V \setminus \{o\} \quad (3.5)$$

and  $\Psi(o) = 1$ . This is an increasing function of  $|x|$  and hence conductances can be defined via

$$c(e) = \frac{1}{\Psi(e^+) - \Psi(e^-)}, \quad (3.6)$$

so the conductance of an edge is inversely proportional to the difference of the gauges at its head and tail vertices. The choice of conductances (3.6) implies for any vertex  $x$ ,

$$\begin{aligned} \Psi(x) &= \Psi(o) + \sum_{o < y \leq x} (\Psi(e(y)^+) - \Psi(e(y)^-)) \\ &= 1 + \sum_{o < y \leq x} \frac{1}{c(e(y))} \\ &= 1 + \frac{1}{\mathcal{C}(o \longleftrightarrow x)} \end{aligned}$$

using the fact that, since we are on a tree, the conductances on edges joining the root to a vertex  $x$  are arranged in series. Hence by definition of  $\Psi$ , we have

$$\frac{1}{\mathbb{P}(o \longleftrightarrow x)} = 1 + \frac{1}{\mathcal{C}(o \longleftrightarrow x)}, \quad (3.7)$$

which is equation (5.12) in [29]. With condition (3.7) satisfied, apply Theorem 5.14 in [29] which states

$$\mathbb{P}(o \longleftrightarrow \infty) \geq \frac{\mathcal{C}(o \longleftrightarrow \infty)}{1 + \mathcal{C}(o \longleftrightarrow \infty)}. \quad (3.8)$$

For the choice of conductances (3.6), let  $i : E \rightarrow [0, \infty)$  denote the unique flow of unit strength with minimum energy, which by Thomson's principle (3.4) satisfies

$$\mathcal{E}(i) = \frac{1}{\mathcal{C}(o \longleftrightarrow \infty)}. \quad (3.9)$$

Then (3.8) reduces to

$$\mathbb{P}(o \longleftrightarrow \infty) \geq \frac{1}{1 + \mathcal{E}(i)}. \quad (3.10)$$

Since by definition of  $i$ ,  $\mathcal{E}(i) \leq \mathcal{E}(\theta)$  for any other unit flow  $\theta$ , it is enough to find any unit flow of finite energy. Also we may normalise any nonzero flow of finite energy by setting  $\theta'(e) = \theta(e) / \sum_{d \in E: e^- = o} \theta(d)$  for each  $e \in E$  to obtain a unit flow  $\theta'$  of finite energy. So we just need to find a nonzero flow of finite energy. To find such a flow, the aim is to use the following result from [13]:

**Proposition 3.2.1** (Proposition 16 in [13]). *For any collection of positive numbers  $(u_e)_{e \in E}$  such that  $\sum_{e:|e|=1} u_e = 1$  and*

$$\inf_{\pi \in \Pi} \sum_{e \in \pi} u_e c(e) > 0,$$

*there exists a non-zero flow whose energy is upper bounded by  $\lim_{n \rightarrow \infty} \max_{e:|e|=n} \sum_{g \leq e} u_g$ .*

### 3.3 Application to Accessibility Percolation on General Trees

We now make a specific choice for the numbers  $(p_n)_n$  introduced in the previous section to enable us to prove Proposition 3.3.1 below. Let  $(f(j))_{j \geq 1} \subseteq [0, 1)$  be a sequence of (non-negative) real numbers. Define  $a_0 = 0$  and for each  $i \in \mathbb{N}$ ,

$$a_i(f) = a_i = \sum_{j=1}^i f(j) \prod_{m=1}^{j-1} (1 - f(m)). \quad (3.11)$$

Given a sequence  $(U_i)_{i \geq 1}$  of i.i.d.  $\text{Unif}([0, 1])$  random variables, we define the events  $\mathcal{A}_i = \mathcal{A}_i(f) = \{U_i \in [a_{i-1}, a_i]\}$ . Observe that

$$\mathbb{P}(\mathcal{A}_i) = a_i - a_{i-1} = f(i) \prod_{m=1}^{i-1} (1 - f(m))$$

for all  $i \in \mathbb{N}$ . Moreover, by construction, for each  $h \in \mathbb{N}$  we have

$$\bigcap_{i=1}^h \mathcal{A}_i \subseteq \{U_1 < \dots < U_h\}$$

and consequently for each  $h \in \mathbb{N}$ ,

$$q_h = \mathbb{P}\left(\bigcap_{i=1}^h \mathcal{A}_i\right) \leq \mathbb{P}(U_1 < \dots < U_h) = \frac{1}{h!}. \quad (3.12)$$

Let  $\mathcal{T}_h$  denote the set of vertices at level  $h$ . Given a path  $(e_1, \dots, e_h)$  of edges from the root  $o$  to a leaf in  $\mathcal{T}_h$  (so that  $e_1^- = o$  and  $e_h^+ \in \mathcal{T}_h$ ), we let  $U_1, \dots, U_h$  denote the independent  $\text{Unif}([0, 1])$  random variables associated to the vertices along the path, which we refer to as *fitness values*, i.e.  $U_1$  is the fitness of the vertex  $e_1^+$ ,  $U_2$  is the fitness of the vertex  $e_2^+$ , and so on. Consider the following percolation process. We say that  $e_i$  is *open* if  $\mathcal{A}_i$  occurs. Since the  $U_i$  are independent and the intervals  $[a_{i-1}, a_i]$  are disjoint, we see that  $q_h = \prod_{j=1}^h f(j) \prod_{m=1}^{j-1} (1 - f(m))$ , the latter being the probability that all the edges  $e_1, \dots, e_h$  are simultaneously open (or, equivalently, the probability that the path of edges  $e_1, \dots, e_h$  is retained in the percolation graph). Because of (3.12), if we can show the existence of an infinite path of open edges from the root to infinity with positive probability, we automatically get (with positive probability) existence of a path from the root to infinity of increasing fitness, i.e. an infinite accessible path.

We now detail the setup needed for Proposition 3.3.1 below. Set  $p_n = \mathbb{P}(\mathcal{A}_n)$  so that  $p_1 = f(1)$  and

$$p_n = f(n) \prod_{i=1}^{n-1} (1 - f(i)) \quad (3.13)$$

for  $n \geq 2$ . Following the strategy highlighted in Section 3.2, we choose the gauge according to (3.5), which means  $\Psi(o) = 1$  and

$$\Psi(x) = \prod_{n=1}^{|x|} \frac{1}{p_n} = \prod_{n=1}^{|x|} \frac{1}{f(n)} \prod_{i=1}^{n-1} \frac{1}{(1 - f(i))} = \frac{1}{q_{|x|}} \quad (3.14)$$

for  $x \in V \setminus \{o\}$ . Note that for any edge  $e \in E$ ,

$$\frac{\Psi(e^-)}{\Psi(e^+)} = \left( \prod_{n=1}^{|e^-|} \frac{1}{p_n} \right) \left( \prod_{n=1}^{|e^+|} \frac{1}{p_n} \right)^{-1} = p_{|e^+|}. \quad (3.15)$$

From (3.6), the corresponding set of conductances  $(c(e))_{e \in E}$  is given by

$$c(e) = \frac{1}{\Psi(e^+) - \Psi(e^-)} = \frac{\left(\frac{1}{\Psi(e^+)}\right)}{\left(1 - \frac{\Psi(e^-)}{\Psi(e^+)}\right)} = \frac{q_{|e^+|}}{1 - p_{|e^+|}} = \frac{\prod_{n=1}^{|e^+|} f(n) \prod_{i=1}^{n-1} (1 - f(i))}{1 - \left(f(|e^+|) \prod_{i=1}^{|e^+|} (1 - f(i))\right)},$$

where we used (3.14) and (3.15). Next we prove the existence result for general trees.

**Proposition 3.3.1.** *Suppose we assign conductances to a general tree with  $\text{br}_f(\mathcal{T}) > 1$  according to the scheme defined above. Then there exists a nonzero flow with finite energy.*

*Proof.* Since  $\text{br}_f(\mathcal{T}) > 1$  we can choose  $\lambda > 1$  such that

$$\inf_{\pi \in \Pi} \sum_{e \in \pi} (|e|!)^{-\lambda} > 0.$$

Now choose  $\alpha \in (0, \lambda - 1)$  and  $\beta \in (0, \lambda - 1 - \alpha)$ . Then let  $f(n) = n^{-(1+\alpha)}$  and  $g(n) = n^{-\beta}$ . Now set

$$u_e = (1 - p_{|e^+|}) \prod_{i=1}^n g(i).$$

Using that  $c(e) = q_{|e^+|}(1 - p_{|e^+|})^{-1}$ , we get for any cutset  $\pi \in \Pi$ ,

$$\begin{aligned} \sum_{e \in \pi} u_e c(e) &= \sum_{e \in \pi} \prod_{i=1}^{|e^+|} p_i \prod_{i=1}^{|e^+|} g(i) \\ &= \sum_{e \in \pi} \prod_{i=1}^{|e^+|} f(i) \prod_{j=1}^{i-1} (1 - f(j)) \prod_{i=1}^{|e^+|} g(i) \\ &= \sum_{e \in \pi} \prod_{i=1}^{|e^+|} f(i) g(i) \prod_{j=1}^{i-1} (1 - f(j)) \\ &= \sum_{e \in \pi} \prod_{i=1}^{|e^+|} f(i) g(i) e^{\sum_{j=1}^{i-1} \log(1 - f(j))}. \end{aligned}$$

For any  $\nu > 0$  we can choose  $N \in \mathbb{N}$  such that if  $j \geq N$  we have  $\log(1 - f(j)) \geq$

$-(1 + \nu)f(j)$ . Hence

$$\sum_{j=1}^{i-1} \log(1 - f(j)) \geq \sum_{j=1}^N \log(1 - f(j)) - (1 + \nu) \sum_{j=N+1}^{\infty} f(j) \geq -S$$

for some finite  $S > 0$ . Thus continuing from earlier,

$$\begin{aligned} \sum_{e \in \pi} u_e c(e) &\geq e^{-S} \sum_{e \in \pi} \prod_{i=1}^{|e^+|} f(i) g(i) = e^{-S} \sum_{e \in \pi} \prod_{i=1}^{|e^+|} \frac{1}{i^{1+\alpha+\beta}} \\ &\geq e^{-S} \sum_{e \in \pi} \prod_{i=1}^{|e^+|} \frac{1}{i^\lambda} = e^{-S} \sum_{e \in \pi} (|e^+|!)^{-\lambda} \\ &\geq e^{-S} \inf_{\pi \in \Pi} \sum_{e \in \pi} (|e^+|!)^{-\lambda} > 0 \end{aligned}$$

using our initial assumption on  $\lambda$  in the last line.  $\square$

We use a standard first moment approach to prove the next result.

**Proposition 3.3.2.** *Suppose  $br_f(\mathcal{T}) < 1$ . Then  $\mathcal{T}$  contains no accessible paths almost surely.*

*Proof.* For a vertex  $v$  at level  $k \in \mathbb{N}$ , let  $u(v) = (o, u_1, \dots, u_k) \in V^k$  denote the unique sequence of vertices joining the root to  $v$ , i.e.  $u_k = v$  and  $u_i$  is adjacent to  $u_{i+1}$  for each  $1 \leq i \leq k - 1$ . Also, for a cutset  $\pi \in \Pi$ , define the random variable  $N_\pi$  as the number of accessible paths from the root to the terminal vertex of an edge in  $\pi$ , explicitly:

$$N_\pi = \sum_{e \in \pi} \mathbb{1}_{\{X(u(e^+)) \in I_{|e|}\}}.$$

For an arbitrary level  $h \in \mathbb{N}$ , recall that we denote by  $\mathcal{T}_h$  the finite subtree of  $\mathcal{T}$  obtained by truncating  $\mathcal{T}$  so that all its terminal vertices are at level  $h$ , and let  $\Pi_h \subseteq \Pi$  consist of those cutsets of  $\mathcal{T}$  containing edges only in  $\mathcal{T}_h$ . Note that any accessible path from the root to level  $h$  must pass through one edge of every cutset  $\pi \in \Pi_h$ . Using Markov's inequality we have

$$\mathbb{P}(E_h) = \mathbb{P}\left(\min_{\pi \in \Pi_h} N_\pi \geq 1\right) \leq \mathbb{E}\left(\min_{\pi \in \Pi_h} N_\pi\right) \leq \min_{\pi \in \Pi_h} \mathbb{E}(N_\pi) = \min_{\pi \in \Pi_h} \sum_{e \in \pi} (|e|)^{-1}. \quad (3.16)$$

Note that for any  $h \in \mathbb{N}$ ,  $\Pi_h \subseteq \Pi_{h+1} \subseteq \Pi$  and the sequence  $(\Pi_h)_h$  exhausts  $\Pi$ . Applying

(3.16) and using this fact, we have

$$\mathbb{P}(E_\infty) = \lim_{h \rightarrow \infty} \mathbb{P}(E_h) \leq \limsup_{h \rightarrow \infty} \min_{\pi \in \Pi_h} \sum_{e \in \pi} (|e|!)^{-1} = \min_{\pi \in \Pi} \sum_{e \in \pi} (|e|!)^{-1} = 0,$$

where the last equality follows from the condition  $\text{br}_f(\mathcal{T}) < 1$ .  $\square$

### 3.4 Spherically Symmetric Trees in the Critical Case

Here we consider a natural choice of scaling for infinite trees, leading to a parametric family of trees with factorial branching number equal to 1 (the indeterminate critical case for the existence of an accessible path for general trees), but which contain almost surely no accessible paths for all parameter values.

Let  $d(h)$  denote the degree of vertices at level  $h \geq 0$ . Assume that  $d(h) = \lfloor \alpha(h+1) \rfloor \vee 1$  (we will henceforth omit the floor function). Let  $N_h$  denote the number of accessible paths from the root to level  $h$ , that is,

$$N_h = \sum_{u \in \mathcal{P}_h} \mathbb{1}_{\{X(u) \in I_h\}}. \quad (3.17)$$

If one can show that  $\mathbb{P}(N_h \geq 1) \rightarrow 0$  as  $h \rightarrow \infty$ , one obtains

$$\mathbb{P}(E_\infty) = \lim_{h \rightarrow \infty} \mathbb{P}(E_h) \leq \lim_{h \rightarrow \infty} \mathbb{P}(N_h \geq 1) = 0.$$

The required limit is shown in the next result for all  $\alpha \in (0, \infty)$ .

**Proposition 3.4.1.** *For  $\alpha > 0$ , let  $\mathcal{T}_\alpha$  denote the spherically symmetric tree such that vertices at level  $h \geq 0$  have degree  $d(h) = \lfloor \alpha(h+1) \rfloor \vee 1$ . Then  $\mathcal{T}_\alpha$  contains no accessible paths, almost surely.*

*Proof.* First consider the case where  $\alpha \in (0, 1)$ . Note that  $\mathcal{P}_h$  has size bounded above by  $\alpha^h h!$  and each path in  $\mathcal{P}_h$  has probability  $\frac{1}{h!}$  of being accessible. Consequently,

$$\mathbb{E}(N_h) \leq \frac{\alpha^h h!}{h!} = \alpha^h.$$

Hence, Markov's inequality implies

$$\mathbb{P}(E_h) = \mathbb{P}(N_h \geq 1) \leq \mathbb{E}(N_h) \leq \alpha^h.$$

This shows that  $\mathbb{P}(E_\infty) = 0$  when  $\alpha < 1$ .

Now we consider the case  $\alpha \geq 1$ . The aim is to effectively iterate the above argument by splitting the paths from the root to a high level into sufficiently many small subintervals. To this end, let  $K = K(h)$  be some large integer (which depends on  $h$ ), which we leave unspecified for the moment. For a constant  $A \geq 1$ , split the set of levels  $[1, K] \cap \mathbb{Z}$  into a fixed number of disjoint subsets  $[1, h] \cap \mathbb{Z}, [h+1, Ah] \cap \mathbb{Z}, [Ah+1, A^2h] \cap \mathbb{Z}, \dots, [A^{H-1}h+1, A^Hh] \cap \mathbb{Z}$  and  $[A^Hh+1, K] \cap \mathbb{Z}$  where  $H \in \mathbb{N}$  is independent of  $h$ . The aim is then to show that for each  $1 \leq i \leq H$  and sufficiently small  $\varepsilon > 0$ , there is no accessible path from the root to a vertex at level  $K$  such that the fitness values of the vertices at levels  $j \in \{A^{i-1}h+1, \dots, A^ih\}$  all lie within the interval  $((i-1)\varepsilon, i\varepsilon]$ . This means that if there is an accessible path, say  $u$ , from the root to a vertex at level  $K$ , then the vertices of  $u$  at levels  $j \in \{A^Hh+1, \dots, K\}$  must all have fitness above  $H\varepsilon$ . By taking  $H > \frac{1-\varepsilon}{\varepsilon}$  we see that  $H\varepsilon > 1 - \varepsilon$ , and so the vertices of  $u$  at levels  $A^Hh+1, \dots, K$  must all have fitness greater than  $1 - \varepsilon$ . By taking  $K$  sufficiently large we expect this probability to be small, giving us the desired conclusion.

To carry out this procedure, define for any  $\varepsilon \in (0, 1)$ ,  $1 \leq i \leq H$  with  $1 - \varepsilon < H\varepsilon \leq 1$ ,  $h \in \mathbb{N}$  and  $A > 1$ , the random variables

$$N_{A^{i-1}h, A^ih}^{((i-1)\varepsilon, i\varepsilon]} = \sum_{u \in \mathcal{P}_{A^ih}} \mathbb{1}_{\{X(u) \in I_{A^ih}, X_j(u) \in ((i-1)\varepsilon, i\varepsilon] \forall j \in [A^{i-1}h, A^ih]\}}. \quad (3.18)$$

In words,  $N_{A^{i-1}h, A^ih}^{((i-1)\varepsilon, i\varepsilon]}$  counts the number of accessible paths to level  $A^ih$  such that the vertices at levels  $A^{i-1}h+1, A^{i-1}h+2, \dots, A^ih$  all have fitness values in the interval  $((i-1)\varepsilon, i\varepsilon]$ . The next step is to compute the expectation of the random variables defined in (3.18). We start by fixing

$$0 < \varepsilon < \min \left\{ \left( \frac{1}{\alpha} \right)^{\frac{A}{A-1}} A^{\frac{1}{A-1}} \left( 1 - \frac{1}{A} \right), 1 \right\}. \quad (3.19)$$

We have

$$\begin{aligned} \mathbb{E} \left( N_{A^{i-1}h, A^ih}^{((i-1)\varepsilon, i\varepsilon]} \right) &= \alpha^{A^ih} (A^ih)! \frac{1}{(A^{i-1}h)!} \frac{\varepsilon^{A^ih - A^{i-1}h}}{(A^ih - A^{i-1}h)!} \\ &= (\alpha\varepsilon)^{A^ih} \left( \frac{1}{\varepsilon} \right)^{A^{i-1}h} \frac{(A^ih)!}{(A^{i-1}h)! (A^ih - A^{i-1}h)!}. \end{aligned} \quad (3.20)$$

Recall Stirling's approximation which implies  $C_1 k^k e^{-k} k^{1/2} \leq k! \leq C_2 k^k e^{-k} k^{1/2}$  for positive constants  $C_1$  and  $C_2$ . Applying the above approximation, we see that the



right-hand side of (3.20) is of order

$$\begin{aligned}
& (\alpha\varepsilon)^{A^i h} \varepsilon^{-A^{i-1} h} \frac{(A^i h)^{A^i h} e^{A^{i-1} h} e^{A^i h - A^{i-1} h}}{e^{A^i h} (A^{i-1} h)^{A^{i-1} h} (A^i h - A^{i-1} h)^{A^i h - A^{i-1} h}} \left( \frac{(A^i h)^{\frac{1}{2}}}{(A^{i-1} h)^{\frac{1}{2}} (A^i h - A^{i-1} h)^{\frac{1}{2}}} \right) \\
&= (\alpha\varepsilon)^{A^i h} \varepsilon^{-A^{i-1} h} \frac{(A^i h)^{A^i h}}{(A^{i-1} h)^{A^{i-1} h} (A^i h)^{A^i h (1 - \frac{1}{A})} \left(1 - \frac{1}{A}\right)^{A^i h (1 - \frac{1}{A})}} \left( \frac{(A^i h)^{\frac{1}{2}}}{A^{i-1} h (A-1)^{\frac{1}{2}}} \right) \\
&= (\alpha\varepsilon)^{A^i h} \varepsilon^{-A^{i-1} h} A^{A^{i-1} h} \left( \frac{1}{\left(1 - \frac{1}{A}\right)^{A-1}} \right)^{A^{i-1} h} \frac{A^{1 - \frac{i}{2}}}{(A-1)^{\frac{1}{2}}} h^{-\frac{1}{2}} \\
&= (\alpha\varepsilon)^{A^i h} \left[ \left( \frac{A}{\varepsilon} \frac{1 - \frac{1}{A}}{\left(1 - \frac{1}{A}\right)^A} \right)^{\frac{1}{A}} \right]^{A^i h} \frac{A^{1 - \frac{i}{2}}}{(A-1)^{\frac{1}{2}}} h^{-\frac{1}{2}} \\
&= \left[ (\alpha\varepsilon) \left( \frac{A-1}{\varepsilon} \right)^{1/A} \frac{A}{A-1} \right]^{A^i h} \frac{A^{1 - \frac{i}{2}}}{(A-1)^{\frac{1}{2}}} h^{-\frac{1}{2}}, \tag{3.21}
\end{aligned}$$

and the last term in square brackets in (3.21) is less than 1 under condition (3.19). Therefore, since  $A^i \geq A$  (as  $A > 1$ ) we obtain that

$$\begin{aligned}
\sum_{i=1}^H \mathbb{P} \left( N_{A^{i-1} h, A^i h}^{((i-1)\varepsilon, i\varepsilon]} \geq 1 \right) &\leq \left( \frac{A}{A-1} \right) \frac{1}{\sqrt{h}} \sum_{i=1}^H \left[ (\alpha\varepsilon) \left( \frac{A-1}{\varepsilon} \right)^{\frac{1}{A}} \frac{A}{A-1} \right]^{A^i h} \\
&\leq H \left( \frac{A}{A-1} \right) \frac{1}{\sqrt{h}} \left[ (\alpha\varepsilon) \left( \frac{A-1}{\varepsilon} \right)^{\frac{1}{A}} \frac{A}{A-1} \right]^{A^h},
\end{aligned}$$

and the term on the right-hand side of the last inequality tends to 0 as  $h \rightarrow \infty$ . Therefore a union bounds yields that, as  $h \rightarrow \infty$ ,

$$\mathbb{P} \left( \exists i \in [1, H] : N_{A^{i-1} h, A^i h}^{((i-1)\varepsilon, i\varepsilon]} \geq 1 \right) \rightarrow 0 \tag{3.22}$$

and hence we can focus our attention on providing an upper bound for

$$\mathbb{P} \left( N_K \geq 1, N_{A^{i-1} h, A^i h}^{((i-1)\varepsilon, i\varepsilon]} = 0 \forall i \in [1, H] \right). \tag{3.23}$$

Observe that, if the event in the above probability were to occur, there would exist at least one increasing subpath between levels  $A^H h + 1$  and  $K = K(h) = A^{H+1} h$  such that the fitness of each vertex in the subpath is greater than  $H\varepsilon$ . Therefore, the probability

in (3.23) is at most

$$\mathbb{P}\left(N_{A^H h, A^{H+1} h}^{(H\varepsilon, 1]} \geq 1\right) \leq \mathbb{E}\left(N_{A^H h, A^{H+1} h}^{(H\varepsilon, 1]}\right).$$

By similar computations to those carried out earlier, it can be seen that the last expectation is of order

$$\left[\alpha(1 - H\varepsilon) \left(\frac{A-1}{1-H\varepsilon}\right)^{\frac{1}{A}} \frac{A}{A-1}\right]^{A^{H+1}h} \left(\frac{A^{1-\frac{H+1}{2}}}{A-1}\right) h^{-\frac{1}{2}}. \quad (3.24)$$

Since  $H > \varepsilon^{-1}(1-\varepsilon)$  we see that  $1-H\varepsilon < \varepsilon$  and hence the above expression is bounded above by

$$\left(\frac{A}{A-1}\right) \frac{1}{\sqrt{h}} \left[\alpha\varepsilon \left(\frac{A-1}{\varepsilon}\right)^{\frac{1}{A}} \frac{A}{A-1}\right]^{A^{H+1}h}.$$

By our condition on  $\varepsilon$  (see (3.19)), the last expression goes to 0 as  $h \rightarrow \infty$  (recall that  $H$  is independent of  $h$ ). Therefore the probability in (3.23) also tends to 0 as  $h \rightarrow \infty$ . Combining this with (3.22) we obtain the desired result, i.e. the probability that there exists an accessible path converges to 0 as  $h \rightarrow \infty$ . Thus we also have that  $\mathbb{P}(E_\infty) = 0$  for  $\alpha \in [1, \infty)$ .  $\square$

### 3.5 Possible Extension to Dynamical Version

In this section, we consider a possible extension of the above theory to the dynamical case, in which we look at the dynamic House of Cards model on an infinite, locally finite tree. This means we associate to each vertex of a tree  $G$  a Poisson process  $x_v(t)$  of intensity  $\lambda_T \otimes \lambda_F$  on  $[0, 1] \times [0, 1]$ , where  $\lambda_T$  and  $\lambda_F$  both denote Lebesgue measure on  $[0, 1]$  (the first of these relates to the time-set and the second to the set of available fitness values). Thus each vertex has its fitness resampled from a uniform distribution on  $[0, 1]$  at the arrival times of independent rate 1 Poisson processes. For  $t \in [0, 1]$  and  $h \in \mathbb{N}$ , we define  $E_{t,h}$  to be the event that there exists an accessible path to level  $h$  at time  $t$ , and define  $E_{t,\infty} = \bigcap_{h>0} E_{t,h}$  to be the event that there exists an accessible path to infinity at time  $t$ . Let  $\partial G$  denote the set of paths in  $G$  without terminal vertices, which we call *rays*. It is well known (e.g. from Lyons and Peres' book [29]) that  $\partial G$  can be made into a (compact) metric space with metric  $d(v, w) = e^{-c(v,w)}$  where  $c(v, w)$  denotes the level of the highest vertex shared by the rays  $v$  and  $w$ .

The study of dynamical Bernoulli percolation on general trees was investigated by

Khoshnevisan [25]. Below we consider how the approach of [25] could be adapted to show the existence of times of accessibility for infinite trees. For any topological space  $A$  and Borel function  $h : A \times A \rightarrow \mathbb{R}$ , the  $h$ -energy of a Borel measure  $\mu \in \mathcal{M}(A)$  is defined as

$$\mathcal{E}_h(\mu) = \iint_{A \times A} h d(\mu \otimes \mu).$$

The  $h$ -capacity of a subset  $F \subseteq A$  is defined as the reciprocal of the infimum of the  $h$ -energy over all Borel probability measures defined on  $F$ :

$$\text{Cap}_h(F) = \frac{1}{\inf_{\mu \in \mathcal{M}_1(F)} \mathcal{E}_h(\mu)}.$$

In our case we set  $A = \partial G \times [0, 1]$  and define the function  $h : A \times A \rightarrow \mathbb{R}$  by

$$h((v, s), (w, t)) = \prod_{i=1}^{c(v,w)} \left( 1 + \left( \frac{1-p_i}{p_i} \right) e^{-|t-s|} \right)$$

Then it follows from basic arguments that

$$\mathbb{P}\left( \bigcup_{t \in [0,1]} E_{t,\infty} \right) \geq \text{Cap}_h(\partial G \times [0, 1]).$$

Define the function  $\psi : \partial G \times \partial G \rightarrow \mathbb{R}$  via

$$\psi(v, w) = \int_0^1 \int_0^1 h((v, s), (w, t)) ds dt = \int_0^1 \int_0^1 \prod_{i=1}^{c(v,w)} \left( 1 + \left( \frac{1-p_i}{p_i} \right) e^{-|t-s|} \right) ds dt.$$

Now observe that if  $\lambda$  denotes Lebesgue measure on  $[0, 1]$  then for any  $F \subseteq \partial G$ ,

$$\inf_{\nu \in \mathcal{M}_1(F \times [0,1])} \mathcal{E}_h(\nu) \leq \inf_{\mu \in \mathcal{M}_1(F)} \mathcal{E}_h(\mu \otimes \lambda) = \inf_{\mu \in \mathcal{M}_1(F)} \mathcal{E}_\psi(\mu).$$

Consequently,

$$\text{Cap}_h(\partial G \times [0, 1]) \geq \text{Cap}_\psi(\partial G).$$

From [29] (see Exercise 15.13) it is known that probability measures on  $\partial G$  are in one-to-one correspondence with unit flows on  $\partial G$ . Also, if there exists a gauge function  $\Psi : V \rightarrow [0, \infty]$  such that  $h(v, w) = \Psi(c(v, w))$ , then defining conductances via  $c(e) = (\Psi(e^+) - \Psi(e^-))^{-1}$  will cause the energy of a flow  $\theta$  with respect to this set of conductances to coincide with the  $\psi$ -energy of  $\mu$ , where  $\mu$  is the probability measure on  $\partial G$  corresponding to  $\theta$ . Thus, if one can show that there exists a unit flow with finite

energy for these conductances, it will immediately follow that there exists a probability measure  $\mu \in \mathcal{M}_1(\partial G)$  such that  $\mathcal{E}_\psi(\mu) < \infty$ , which will imply  $\text{Cap}_\psi(\partial G) > 0$ . To get an explicit lower bound on  $\text{Cap}_\psi(\partial G)$ , consider the function

$$R(n) = \int_0^1 \int_0^1 \prod_{i=1}^n \left( 1 + \left( \frac{1-p_i}{p_i} \right) e^{-|t-s|} \right) ds dt = 2 \int_0^1 \int_0^u \prod_{i=1}^n \left( 1 + \left( \frac{1-p_i}{p_i} \right) e^{-v} \right) dv du$$

for  $n \geq 1$ . Similarly to [25], the aim here is to obtain an upper bound for  $R(n)$  of the form  $Q(n) \prod_{i=1}^n p_i^{-1}$ , where  $Q(n)$  decays to zero rapidly with  $n$ . It will then be sufficient to show that  $\text{Cap}_g(\partial G) > 0$  where for rays  $v, w \in \partial G$  we define

$$g(v, w) = Q(c(v, w)) \prod_{i=1}^{c(v, w)} \frac{1}{p_i}.$$

To this end, define for vertices  $x \in V(G) \setminus \{o\}$  the gauge

$$\Psi(x) = Q(|x|) \prod_{i=1}^{|x|} \frac{1}{p_i}$$

and set  $\Psi(o) = 0$ . Conductances can then be generated in the usual way:

$$\begin{aligned} c(e) &= \frac{1}{\Psi(e^+) - \Psi(e^-)} \\ &= \left( Q(|e|) \prod_{i=1}^{|e|} \frac{1}{p_i} - Q(|e| - 1) \prod_{i=1}^{|e|-1} \frac{1}{p_i} \right)^{-1} \\ &\approx \frac{1}{Q(|e|)} \prod_{i=1}^{|e|} p_i. \end{aligned}$$

Provided  $Q(n)$  decays sufficiently rapidly, it may then be possible (for spherically symmetric trees) to find a criterion in terms of the degree sequence of the tree such that for these conductances, there exists a unit flow of finite energy. If this criterion covers trees known to lack accessible paths in the static case, one could deduce the existence of exceptional times of accessibility.

## Chapter 4

# Invasion in a Periodic Environment

### 4.1 Model

Suppose we begin with  $N_1(0) = 1$  and  $N_0(0) = \lfloor \bar{n}_0 K \rfloor$  where  $\bar{n}_0$  is the (Lotka-Volterra) positive equilibrium density of the type 0 population given by

$$\bar{n}_0 = \frac{b_0 - d_0}{C_{0,0}}.$$

Define the function  $b_1(t)$  via

$$b_1(t) = \begin{cases} b_1^+ & \text{for } t \in \bigcup_{m=0}^{\infty} [2mS_K, (2m+1)S_K] \\ b_1^- & \text{for } t \in \bigcup_{m=0}^{\infty} [(2m+1)S_K, 2(m+1)S_K) \end{cases}$$

where  $b_1^- < b_1^+$  and the period  $S_K \in (0, \infty)$  is allowed to depend on the carrying capacity  $K \in \mathbb{N}$ . Also let  $d_1(t) \geq 0$  be a right-continuous function acting as the intensity function of the Poisson point process recording the death events of type 1 individuals. For the main theorem in this section (Theorem 4.3.2) we will additionally assume that  $\sup_{t \geq 0} d_1(t) = d_1 < b_1^-$  which means in the absence of competition ( $C_{0,1} = C_{1,1} = 0$ ) the population size of the type 1 individuals follows a (time varying) supercritical birth-death process. This assumption will be weakened for some of the results in Section 4.3. Let  $Q_0$  and  $Q_1$  be independent Poisson random measures of unit intensity on  $[0, \infty)^2$ .

Then the population processes of the type 0 and type 1 individuals are given by

$$N_0(t) = \int_0^\infty \int_0^t \left( \mathbb{1}_{\{0 \leq \theta \leq b_0 N_0(s)\}} - \mathbb{1}_{\{b_0 N_0(s) \leq \theta \leq (b_0 + d_0 + C_{0,0} N_0(s) + C_{0,1} N_1(s)) N_0(s)\}} \right) Q_0(ds, d\theta) + \lfloor \bar{n}_0 K \rfloor$$

and

$$N_1(t) = \int_0^\infty \int_0^t \left( \mathbb{1}_{\{0 \leq \theta \leq b_1(s) N_1(s)\}} - \mathbb{1}_{\{b_1(s) N_1(s) \leq \theta \leq (b_1(s) + d_1(s) + C_{1,1} N_1(s) + C_{0,1} N_0(s)) N_1(s)\}} \right) Q_1(ds, d\theta) + 1$$

respectively. These are time-varying logistic birth-death processes. For  $i \in \{0, 1\}$  and  $R \in \mathbb{R}$  introduce the stopping times

$$T_R^{(i)} = \{t \geq 0 : N_i(t) \geq R\}.$$

Also introduce the interval

$$I_\varepsilon^{(0)} = \left[ K \left( \bar{n}_0 - \frac{\varepsilon C_{1,0}}{C_{0,0}} \right), K \left( \bar{n}_0 + \frac{\varepsilon C_{1,0}}{C_{0,0}} \right) \right]$$

which is roughly an  $\varepsilon K$ -neighbourhood of the type 0 equilibrium size (depending on the ratio of competition pressures), and the stopping time

$$\tilde{T}_\varepsilon^{(0)} = \inf \{t \geq 0 : N_0(t) \notin I_\varepsilon^{(0)}\}.$$

Now define the adjusted death rate functions (motivated by the set  $I_\varepsilon^{(0)}$  defined above)

$$d_1^-(t) = d_1(t) + C_{1,0} \left( \bar{n}_0 + \frac{\varepsilon C_{1,0}}{C_{0,0}} \right) + \varepsilon C_{1,1}, \quad d_1^+(t) = d_1(t) + C_{1,0} \left( \bar{n}_0 - \frac{\varepsilon C_{1,0}}{C_{0,0}} \right)$$

so that for each  $* \in \{-, +\}$ , we can define an approximated population size process for the type 1 individuals, namely

$$N_1^*(t) = 1 + \int_0^\infty \int_0^t \left( \mathbb{1}_{\{0 \leq \theta \leq b_1(s) N_1^*(s)\}} - \mathbb{1}_{\{b_1(s) N_1^*(s) \leq \theta \leq (b_1(s) + d_1^*(s)) N_1^*(s)\}} \right) Q_1(ds, d\theta).$$

which is a standard time-varying birth-death process (as the quadratic death term is replaced by a linear death term). By our choices of  $d_-(t)$  and  $d_+(t)$  as well as the fact

that  $N_1^-(t)$  and  $N_1^+(t)$  are defined on the same probability space as  $N_1(t)$ , the event

$$\{N_1^-(t) \leq N_1(t) \leq N_1^+(t) \text{ for all } t \in [0, \tilde{T}_\varepsilon^{(0)}]\}$$

holds almost surely. Note that in the particular case that the intrinsic death intensity function  $d_1(t)$  takes a constant value in odd periods of length  $S_K$  and a different, smaller value in even periods of length  $S_K$ , we define  $d_1^{*,+}$  such that  $d_1^*(t) = d_1^{*,+}$  in odd periods and  $d_1^-(t) = d_1^{*, -}$  in even periods. To approximate the type 0 process, for each  $* \in \{-, +\}$  we can define

$$N_0^*(t) = \int_0^\infty \int_0^t \left( \mathbb{1}_{\{0 \leq \theta \leq b_0(s)N_0^*(s)\}} - \mathbb{1}_{\{b_0(s)N_0^*(s) \leq \theta \leq (b_0(s) + d_0^*(s) + C_{0,0}N_0^*(s))N_0^*(s)\}} \right) Q_0(ds, d\theta) + \lfloor \bar{n}_0 K \rfloor$$

where

$$d_0^-(t) = d_0(t) + \varepsilon C_{0,1} \text{ and } d_0^+(t) = d_0(t).$$

As for the type 1 approximating processes, the event

$$\{N_0^-(t) \leq N_0(t) \leq N_0^+(t) \text{ for all } t \in [0, T_\varepsilon^{(1)}]\}$$

holds almost surely.

For some of the results in the next sections we will consider single non-density dependent periodic birth-death processes with more general birth and death intensities  $b_K(s)$  and  $d_K(s)$  which we assume to be bounded, non-negative functions (uniformly in  $K$ ). We also assume  $b_K(t)$  and  $d_K(t)$  are periodic functions with coinciding periods of length  $2S_K$ . In particular, letting  $Q$  be an independent Poisson random measure of unit intensity on  $[0, \infty)^2$ , we define the process  $N_K(t)$  according to

$$N_K(t) = 1 + \int_0^\infty \int_0^t \left( \mathbb{1}_{\{0 \leq \theta \leq b_K(s)N_K(s)\}} - \mathbb{1}_{\{b_K(s)N_K(s) \leq \theta \leq (b_K(s) + d_K(s))N_K(s)\}} \right) Q(ds, d\theta) \quad (4.1)$$

In Section 4.4 we will also consider two coupled processes  $(N_0(t), N_1(t))$  such that the process  $N_1(t)$  (which we think of as the size of the invading population) has  $2S_K$ -periodic intrinsic birth and death intensities  $b_K(t)$  and  $d_K(t)$ , and otherwise the system behaves like coupled logistic birth-death processes with competition parameters  $C_{0,0}$ ,

$C_{0,1}$  and  $C_{1,1}$ . In particular, given two independent Poisson processes of rate 1,  $P_b(t)$  and  $P_d(t)$ , we use the representation

$$\begin{aligned}
N_0(t) &= N_0(0) + P_b\left(b_0 \int_0^t N_0(s) ds\right) \\
&\quad - P_d\left(\int_0^t \left(d_0 + \frac{C_{0,0}N_0(s)}{K} + \frac{C_{0,1}N_1(s)}{K}\right) N_0(s) ds\right) \\
N_1(t) &= N_1(0) + P_b\left(\int_0^t b_K(s) N_1(s) ds\right) \\
&\quad - P_d\left(\int_0^t \left(d_K(s) + \frac{C_{0,1}N_0(s)}{K} + \frac{C_{1,1}N_1(s)}{K}\right) N_1(s) ds\right).
\end{aligned} \tag{4.2}$$

We will also need some existing results on standard birth-death processes. The first part of the next Lemma comes from part 1 of Lemma A.1 in [5] and the second part comes from part 3 of Lemma A.1 in [5].

**Lemma 4.1.1** (Lemma A.1 in [5]).

- (i) Let  $N(t)$  denote a standard birth-death process of birth rate  $b$  and death rate  $d$  with initial condition  $N(0) = j$  for some  $j \in \mathbb{N}$ . For  $R \in [0, \infty)$  let  $T_R = \inf\{t \geq 0 : N(t) = \lfloor R \rfloor\}$ . Then for any positive integers  $i < j$  and  $k > j$ ,

$$\mathbb{P}(T_k < T_i) = \frac{1 - \left(\frac{d}{b}\right)^{j-i}}{1 - \left(\frac{d}{b}\right)^{k-i}}.$$

- (ii) Let  $N(t)$  denote a standard birth-death process of birth rate  $b$  and death rate  $d < b$  with initial condition  $N(0) = 1$ . For  $R \in [0, \infty)$  let  $T_R = \inf\{t \geq 0 : N(t) = \lfloor R \rfloor\}$ . Then for any  $\eta > 0$ ,

$$\mathbb{P}\left(\left|\frac{T_N}{\log N} - \frac{1}{b-d}\right| > \eta \mid T_0 = \infty\right) \rightarrow 0$$

as  $N \rightarrow \infty$ .

The following is a slightly rephrased version of Lemma A.2 of [10]:

**Lemma 4.1.2** (Lemma A.2 of [10]). Let  $N(t)$  denote a standard birth-death process of birth rate  $b$  and death rate  $d < b$  with initial condition  $N(0) = K^\gamma$  for some  $\gamma > 0$ . Then for any  $\eta > 0$  and  $T > 0$ ,

$$\mathbb{P}\left(\sup_{t \in [0, T]} \left|\frac{\log N(t \log K)}{\log K} - (\gamma + (b-d)t)\right| > \eta\right) \rightarrow 0$$



as  $K \rightarrow \infty$ .

By rearranging the above result we have the following corollary:

**Corollary 4.1.3.** *Let  $N(t)$  denote a standard birth-death process of birth rate  $b$  and death rate  $d < b$  with initial condition  $N(0) = K^\gamma$  for some  $\gamma > 0$ . For  $R \in [0, \infty)$  let  $T_R = \inf \{t \geq 0 : N(t) = \lfloor R \rfloor\}$ . Then for any  $\beta > \gamma$  and  $\eta > 0$  we have*

$$\mathbb{P}\left(\left|\frac{T_{K^\beta}}{\log K} - \frac{\beta - \gamma}{b - d}\right| > \eta\right) \rightarrow 0 \quad (4.3)$$

as  $K \rightarrow \infty$ .

*Proof.* Set  $\eta' \in [0, \eta(b - d))$ . Note that

$$\begin{aligned} \mathbb{P}\left(T_{K^\beta} < \left(\frac{\beta - \gamma}{b - d} - \eta\right) \log(K)\right) &\leq \mathbb{P}\left(\sup_{t \in [0, \frac{\beta - \gamma}{b - d} - \eta]} \frac{\log N(t \log K)}{\log K} \geq \beta\right) \\ &\leq \mathbb{P}\left(\sup_{t \in [0, \frac{\beta - \gamma}{b - d} - \eta]} \left(\frac{\log N(t \log K)}{\log K} - (\gamma + (b - d)t)\right) \geq \beta - \sup_{t \in [0, \frac{\beta - \gamma}{b - d} - \eta]} (\gamma + (b - d)t)\right) \\ &\leq \mathbb{P}\left(\sup_{t \in [0, \frac{\beta - \gamma}{b - d} - \eta]} \left|\frac{\log N(t \log K)}{\log K} - (\gamma + (b - d)t)\right| > \eta'\right) \rightarrow 0 \end{aligned}$$

where the convergence to zero is implied by Lemma 4.1.2. Similarly we have

$$\begin{aligned} \mathbb{P}\left(T_{K^\beta} > \left(\frac{\beta - \gamma}{b - d} + \eta\right) \log(K)\right) &= \mathbb{P}\left(\sup_{t \in [0, \frac{\beta - \gamma}{b - d} + \eta]} \frac{\log N(t \log K)}{\log K} < \beta\right) \\ &\leq \mathbb{P}\left(\sup_{t \in [0, \frac{\beta - \gamma}{b - d} + \eta]} \left((\gamma + (b - d)t) - \frac{\log N(t \log K)}{\log K}\right) > \sup_{t \in [0, \frac{\beta - \gamma}{b - d} + \eta]} (\gamma + (b - d)t) - \beta\right) \\ &\leq \mathbb{P}\left(\sup_{t \in [0, \frac{\beta - \gamma}{b - d} + \eta]} \left|\frac{\log N(t \log K)}{\log K} - (\gamma + (b - d)t)\right| > \eta'\right) \rightarrow 0 \end{aligned}$$

using Lemma 4.1.2 in the last line. Combining with the result of the previous display gives (4.3).  $\square$

We will also use the following result, which deals with logistic birth-death processes and comes from part 2 of Lemma A.2 in [5].

**Lemma 4.1.4** (Lemma A.2 in [5]). *For a fixed birth rate  $b > 0$ , death rate  $d > 0$  and competition parameter  $C > 0$ , let  $\bar{n} = (b - d)/C$  denote the nonzero equilibrium density of the Lotka-Volterra system*

$$\frac{dn}{dt} = (b - d)n - Cn^2,$$

given  $n(0) > 0$ . For  $K \in \mathbb{N}$ , let  $N^K(t)$  denote the logistic birth-death process of the form

$$N^K(t) = \lfloor \bar{n}K \rfloor + \int_0^\infty \int_0^t \left( \mathbb{1}_{\{0 \leq \theta \leq bN^K(s)\}} - \mathbb{1}_{\{bN^K(s) \leq \theta \leq (b+d)N^K(s)\}} \right) Q(ds, d\theta)$$

for a given unit-intensity Poisson random measure  $Q$  on  $[0, \infty)^2$ . Then for every  $\varepsilon > 0$  there exists a constant  $V > 0$  such that

$$\mathbb{P} \left( \sup_{t \in [0, e^{VK}]} \left| \frac{N^K(t)}{K} - \bar{n} \right| \leq \varepsilon \right) \rightarrow 1$$

as  $K \rightarrow \infty$ .

## 4.2 Time to Reach a Level

In this section, the aim is to obtain asymptotic estimates on the time taken for the periodic birth-death processes  $N_1^*(t)$ ,  $*$   $\in \{-, +\}$  to reach size  $\varepsilon K$  as  $K \rightarrow \infty$ . The next lemma will be proved for the more general process as described by (4.1). The condition (4.4) essentially requires that the average growth rate of the process can be uniformly bounded below by a positive constant.

**Lemma 4.2.1.** *Define the function  $\rho_Y^K(t) = \int_0^t (d_Y^K(s) - b_Y^K(s)) ds$  and suppose that  $\sup_{K \geq 1} \sup_{t \geq 0} b_K(t) \leq b_{\max}$  for some  $b_{\max} > 0$ . Assume  $\rho_Y^K(t) < 0$  for all  $t \geq 0$  and there exists  $\nu > 0$  such that*

$$\limsup_{K \rightarrow \infty} \sup_{t \geq 0} \left\{ \frac{\rho_K(t)}{t} \right\} < -\nu. \quad (4.4)$$

Let  $(Y_K(t))_{t \geq 0}$  be an inhomogeneous Markov process with the same family of transition operators as  $N_K(t)$ , as determined by (4.1), but with  $Y_K(0) = K^\delta$  for some  $\delta > 0$ . Then for all  $\delta/2 < \eta < \delta$ ,

$$\mathbb{P} \left( \sup_{t \in [0, T \log K]} |e^{\rho_Y^K(t)} Y_K(t) - K^\delta| \geq K^\eta \right) \rightarrow 0$$

as  $K \rightarrow \infty$ .

*Proof.* Note by Lemma 4.6.1 (see Auxiliary results) that  $W_K(t) = e^{\rho_Y^K(t)} Y_K(t)$  is a square integrable martingale, where

$$\rho_Y^K(t) = \int_0^t (d_Y^K(s) - b_Y^K(s)) ds.$$

Additionally the process

$$H_K(t) = Y_K(t) - K^\delta - \int_0^t (b_Y^K(s) - d_Y^K(s)) Y_K(s) ds$$

is a square integrable martingale. The predictable quadratic variation process of  $(H_K(t))_{t \geq 0}$  can be calculated by standard techniques (see Appendix, Section A.1) to give:

$$\langle H_K \rangle(t) = \int_0^t (d_Y^K(s) + b_Y^K(s)) Y_K(s) ds.$$

Now we use Itô's formula (see Appendix, Section A.2) to obtain a representation of  $W_K(t)$  in terms of a martingale:

$$W_K(t) = K^\delta + \hat{H}_K(t)$$

where

$$\hat{H}_K(t) = \int_0^t e^{\rho_Y^K(s)} dH_K(s).$$

Note that  $\hat{H}_K(t)$  is a square integrable martingale and has predictable quadratic variation:

$$\begin{aligned} \langle \hat{H}_K \rangle(t) &= \left\langle \int_0^t e^{\rho_Y^K(s)} dH_K(s) \right\rangle = \int_0^t e^{2\rho_Y^K(s)} d\langle H_K \rangle(s) \\ &= \int_0^t e^{2\rho_Y^K(s)} (d_Y^K(s) + b_Y^K(s)) Y_K(s) ds. \end{aligned}$$

We have for any  $T > 0$  and  $\eta \in (0, \delta)$ ,

$$\begin{aligned} \mathbb{P}\left(\sup_{t \in [0, T \log K]} |e^{\rho_Y^K(t)} Y_K(t) - K^\delta| \geq K^\eta\right) &\leq \mathbb{P}\left(\sup_{t \in [0, T \log K]} |\hat{H}_K(t)| \geq K^\eta\right) \\ &\leq 4K^{-2\eta} \mathbb{E}\left(\langle \hat{H}_K \rangle(T \log K)\right) \\ &= 4K^{-2\eta} \int_0^{T \log K} e^{2\rho_Y^K(s)} (d_Y^K(s) + b_Y^K(s)) K^\delta e^{-\rho_Y^K(s)} ds \\ &= 4K^{\delta-2\eta} \int_0^{T \log K} (d_Y^K(s) + b_Y^K(s)) e^{\rho_Y^K(s)} ds \end{aligned}$$

Consider the integral  $\int_0^{T \log K} (d_Y^K(s) + b_Y^K(s)) e^{\rho_Y^K(s)} ds$ . Setting  $t_K = T \log K$  we have

$$\begin{aligned} \int_0^{t_K} (d_Y^K(s) + b_Y^K(s)) e^{\rho_Y^K(s)} ds &= \int_0^{t_K} (d_Y^K(s) - b_Y^K(s)) e^{\rho_Y^K(s)} ds + 2 \int_0^{t_K} b_Y^K(s) e^{\rho_Y^K(s)} ds \\ &= e^{\rho_Y^K(t_K)} + 2 \int_0^{t_K} b_Y^K(s) e^{\rho_Y^K(s)} ds \\ &\leq e^{\rho_Y^K(t_K)} + 2b_{\max} \int_0^{t_K} e^{\rho_Y^K(s)} ds. \end{aligned}$$

Choose  $K$  large enough so that  $\sup_{s \geq 0} \rho_Y^K(s) < -\nu$  for some  $\nu > 0$ . Then for sufficiently large  $K$ ,

$$\int_0^{t_K} e^{\rho_Y^K(s)} ds \leq \int_0^{t_K} e^{-\nu s} ds = \frac{1 - e^{-\nu t_K}}{\nu}$$

and

$$e^{\rho_Y^K(t_K)} = e^{\frac{\rho_Y^K(t_K)}{t_K} t_K} \leq e^{\left(\sup_{s \geq 0} \frac{\rho_Y^K(s)}{s}\right) t_K} \leq e^{-\nu t_K} = K^{-\nu T}.$$

Thus,

$$\begin{aligned} \mathbb{P}\left(\sup_{t \in [0, T \log K]} |e^{\rho_Y^K(t)} Y_K(t) - K^\delta| \geq K^\eta\right) &\leq 4K^{\delta-2\eta} \int_0^{t_K} (d_Y^K(s) + b_Y^K(s)) e^{\rho_Y^K(s)} ds \\ &\leq 4K^{\delta-2\eta} \left(K^{-\nu T} + \frac{2b_{\max}}{\nu}\right) \\ &\rightarrow 0 \end{aligned}$$

as  $K \rightarrow \infty$ . □

We now specialise to the case where the birth and death intensity functions,  $b_K(t)$  and  $d_K(t)$ , each jump between two values and remain constant for intervals of length  $S_K$  (and hence are periodic functions with period  $2S_K$ ). In particular, we focus on the approximating non-logistic birth-death processes of the type 1 population  $N_1^*(t)$ ,  $\in \{-, +\}$  defined in Section 4.1. We assume that for each  $* \in \{-, +\}^1$  that  $\frac{1}{2}((b_1^+ - d_1^{*,+}) + (b_1^- - d_1^{*,+})) > 0$  and that the process  $N_1^*(t)$  begins in a period with growth rate  $b_1^+ - d_1^{*,+} > 0$ . These conditions ensure the hypotheses of Lemma 4.2.1 above are satisfied. For the next lemma we additionally suppose that for each  $* \in \{-, +\}$ ,  $d_1^*(t) = d_1^*$ , i.e. the death rate stays constant.

**Lemma 4.2.2.** *Let the period  $S_K$  be such that  $S_K \rightarrow \infty$  and  $S_K/\log K \rightarrow 0$  as  $K \rightarrow \infty$ . Suppose also that for  $* \in \{-, +\}$ ,  $d_1^*(t) = d_1^*$ , a constant, for all  $t \geq 0$ . Then*

---

<sup>1</sup>It is worth nothing that  $* = -$  refers to the death rate adjusted for increased competition as defined in the model setup rather than the lower of the two intrinsic death rate values taken in each period.

for any  $\eta > 0$ ,

$$\mathbb{P}\left(\left|\frac{T_{\varepsilon K}^{(1,*)}}{\log(\lceil \varepsilon K \rceil)} - \frac{1}{\frac{1}{2}(b_1^- + b_1^+) - d_1^*}\right| > \eta \mid T_0^{(1,*)} = \infty\right) \rightarrow 0$$

as  $K \rightarrow \infty$ .

*Proof.* For each  $* \in \{-, +\}$  and any  $\delta > 0$ , let  $T_{(\varepsilon K)^\delta}^{(1,*)}$  be the time taken for the process  $N_1^*(t)$  to reach size  $(\varepsilon K)^\delta$  and let  $T_{\varepsilon K, \delta}^{(1,*)}$  be the time taken for the process  $N_1^*(t)$  to reach level  $\varepsilon K$  conditioned on  $N_1^*(0) = (\varepsilon K)^\delta$ . First given  $\eta > 0$ , choose  $\delta > 0$  such that  $\delta < \eta/4$ ,

$$\left|\frac{\delta}{b_1^- - d_1^*} - \frac{\delta}{\frac{1}{2}(b_1^- + b_1^+) - d_1^*}\right| = \delta \left(\frac{1}{b_1^- - d_1^*} - \frac{1}{\frac{1}{2}(b_1^- + b_1^+) - d_1^*}\right) < \frac{\eta}{4}$$

and

$$\left|\frac{\delta}{b_1^+ - d_1^*} - \frac{\delta}{\frac{1}{2}(b_1^- + b_1^+) - d_1^*}\right| = \delta \left(\frac{1}{\frac{1}{2}(b_1^- + b_1^+) - d_1^*} - \frac{1}{b_1^+ - d_1^*}\right) < \frac{\eta}{4}.$$

These conditions will be required later. Then define the event

$$E_K = \left\{ \left| \frac{T_{(\varepsilon K)^\delta}^{(1,*)}}{\log(\lceil \varepsilon K \rceil)} - \frac{\delta}{\frac{1}{2}(b_1^- + b_1^+) - d_1^*} \right| \leq \frac{\eta}{2} \right\}.$$

Now we have

$$\begin{aligned}
& \mathbb{P}\left(\left|\frac{T_{\varepsilon K}^{(1,*)}}{\log(\lceil \varepsilon K \rceil)} - \frac{1}{\frac{1}{2}(b_1^- + b_1^+) - d_1^*}\right| > \eta \mid T_0^{(1,*)} = \infty\right) \\
& \leq \mathbb{P}\left(\left|\frac{T_{\varepsilon K}^{(1,*)}}{\log(\lceil \varepsilon K \rceil)} - \frac{1}{\frac{1}{2}(b_1^- + b_1^+) - d_1^*}\right| > \eta \mid T_0^{(1,*)} = \infty, E_K\right) + \mathbb{P}(E_K^c \mid T_0^{(1,*)} = \infty) \\
& = \mathbb{P}\left(\left|\frac{T_{(\varepsilon K)^\delta}^{(1,*)} + T_{\varepsilon K, \delta}^{(1,*)}}{\log(\lceil \varepsilon K \rceil)} - \frac{1}{\frac{1}{2}(b_1^- + b_1^+) - d_1^*}\right| > \eta \mid T_0^{(1,*)} = \infty, E_K\right) + \mathbb{P}(E_K^c \mid T_0^{(1,*)} = \infty) \\
& \leq \mathbb{P}\left(\left|\frac{T_{\varepsilon K, \delta}^{(1,*)}}{\log(\lceil \varepsilon K \rceil)} - \frac{1 - \delta}{\frac{1}{2}(b_1^- + b_1^+) - d_1^*}\right| > \frac{\eta}{2} \mid T_0^{(1,*)} = \infty, E_K\right) \\
& + \mathbb{P}\left(\left|\frac{T_{(\varepsilon K)^\delta}^{(1,*)}}{\log(\lceil \varepsilon K \rceil)} - \frac{\delta}{\frac{1}{2}(b_1^- + b_1^+) - d_1^*}\right| > \frac{\eta}{2} \mid T_0^{(1,*)} = \infty, E_K\right) + \mathbb{P}(E_K^c \mid T_0^{(1,*)} = \infty) \\
& = \mathbb{P}\left(\left|\frac{T_{\varepsilon K, \delta}^{(1,*)}}{\log(\lceil \varepsilon K \rceil)} - \frac{1 - \delta}{\frac{1}{2}(b_1^- + b_1^+) - d_1^*}\right| > \frac{\eta}{2} \mid T_0^{(1,*)} = \infty, E_K\right) \\
& + \mathbb{P}\left(E_K^c \mid T_0^{(1,*)} = \infty, E_K\right) + \mathbb{P}(E_K^c \mid T_0^{(1,*)} = \infty) \\
& = \mathbb{P}\left(\left|\frac{T_{\varepsilon K, \delta}^{(1,*)}}{\log(\lceil \varepsilon K \rceil)} - \frac{1 - \delta}{\frac{1}{2}(b_1^- + b_1^+) - d_1^*}\right| > \frac{\eta}{2} \mid T_0^{(1,*)} = \infty, E_K\right) + \mathbb{P}(E_K^c \mid T_0^{(1,*)} = \infty)
\end{aligned}$$

To find the limit of  $\mathbb{P}(E_K^c \mid T_0^{(1,*)} = \infty)$ , note that

$$\begin{aligned}
\mathbb{P}(E_K^c \mid T_0^{(1,*)} = \infty) & = \mathbb{P}\left(\left|\frac{T_{(\varepsilon K)^\delta}^{(1,*)}}{\log(\lceil \varepsilon K \rceil)} - \frac{\delta}{\frac{1}{2}(b_1^- + b_1^+) - d_1^*}\right| > \frac{\eta}{2} \mid T_0^{(1,*)} = \infty\right) \\
& = \mathbb{P}\left(\frac{T_{(\varepsilon K)^\delta}^{(1,*)}}{\log(\lceil \varepsilon K \rceil)} < \frac{\delta}{\frac{1}{2}(b_1^- + b_1^+) - d_1^*} - \frac{\eta}{2} \mid T_0^{(1,*)} = \infty\right) \\
& + \mathbb{P}\left(\frac{T_{(\varepsilon K)^\delta}^{(1,*)}}{\log(\lceil \varepsilon K \rceil)} > \frac{\delta}{\frac{1}{2}(b_1^- + b_1^+) - d_1^*} + \frac{\eta}{2} \mid T_0^{(1,*)} = \infty\right)
\end{aligned}$$

Consider the first of the probabilities in the above sum. Define for  $\star \in \{-, +\}$ , the process  $N_1^{*,\star}(t)$  as

$$N_1^{*,\star}(t) = 1 + \int_0^\infty \int_0^t \left( \mathbb{1}_{\{0 \leq \theta \leq b_1^* N_1^{*,\star}(s)\}} - \mathbb{1}_{\{b_1^* N_1^{*,\star}(s) \leq \theta \leq (b_1^* + d_1^*) N_1^{*,\star}(s)\}} \right) Q_1(ds, d\theta),$$

using the same Poisson random measure  $Q_1(ds, d\theta)$  as the process  $N_1^*(t)$ . Note that  $N_1^{*,\star}(t)$  is a standard birth-death process with fixed birth rate  $b_1^*$  and fixed death rate

$d_1^*$ . Also, since  $b_1^- \leq b_1(t) \leq b_1^+$ , defining  $\tau_{(\varepsilon K)^\delta}^{1,*,*} = \inf\{t \geq 0 : N_1^{*,*}(t) \geq (\varepsilon K)^\delta\}$  we have that  $\tau_{(\varepsilon K)^\delta}^{1,*,+} \leq T_{(\varepsilon K)^\delta}^{(1,*)}$  almost surely. Let  $\tau_0^{1,*,*}$  denote the extinction time of the process  $N_1^{*,*}(t)$  and note that  $\mathbb{P}(T_0^{1,*} = \infty) \geq \mathbb{P}(\tau_0^{1,*,*} = \infty) = 1 - (d_*/b_-) > 0$  since  $b_- > d_*$ . Hence,

$$\begin{aligned} & \mathbb{P}\left(\frac{T_{(\varepsilon K)^\delta}^{(1,*)}}{\log(\lceil \varepsilon K \rceil)} < \frac{\delta}{\frac{1}{2}(b_1^- + b_1^+) - d_1^*} - \frac{\eta}{2} \mid T_0^{(1,*)} = \infty\right) \\ & \leq \left(\frac{\mathbb{P}(\tau_0^{1,*,+} = \infty)}{\mathbb{P}(T_0^{1,*} = \infty)}\right) \mathbb{P}\left(\frac{\tau_{(\varepsilon K)^\delta}^{(1,*,+)}}{\log(\lceil \varepsilon K \rceil)} < \frac{\delta}{\frac{1}{2}(b_1^- + b_1^+) - d_1^*} - \frac{\eta}{2} \mid \tau_0^{(1,*,+)} = \infty\right) \\ & \leq R_1 \mathbb{P}\left(\frac{\tau_{(\varepsilon K)^\delta}^{(1,*,+)}}{\log(\lceil \varepsilon K \rceil)} < \frac{\delta}{\frac{1}{2}(b_1^- + b_1^+) - d_1^*} - \frac{\eta}{2} \mid \tau_0^{(1,*,+)} = \infty\right) \end{aligned}$$

for some finite constant  $R_1 > 0$ . Now by our choice of  $\delta$ , we have

$$\begin{aligned} & \mathbb{P}\left(\frac{\tau_{(\varepsilon K)^\delta}^{(1,*,+)}}{\log(\lceil \varepsilon K \rceil)} < \frac{\delta}{\frac{1}{2}(b_1^- + b_1^+) - d_1^*} - \frac{\eta}{2} \mid \tau_0^{(1,*,+)} = \infty\right) \\ & \leq \mathbb{P}\left(\frac{\tau_{(\varepsilon K)^\delta}^{(1,*,+)}}{\log(\lceil \varepsilon K \rceil)} < \frac{\delta}{b_1^+ - d_1^*} - \frac{\eta}{4} \mid \tau_0^{(1,*,+)} = \infty\right) \\ & \rightarrow 0 \end{aligned}$$

using part 2 of Lemma 4.1.1. Again by our choice of  $\delta$ , we have

$$\begin{aligned} & \mathbb{P}\left(\frac{T_{(\varepsilon K)^\delta}^{(1,*)}}{\log(\lceil \varepsilon K \rceil)} < \frac{\delta}{\frac{1}{2}(b_1^- + b_1^+) - d_1^*} - \frac{\eta}{2} \mid T_0^{(1,*)} = \infty\right) \\ & \leq \mathbb{P}\left(\frac{T_{(\varepsilon K)^\delta}^{(1,*)}}{\log(\lceil \varepsilon K \rceil)} > \frac{\delta}{b_1^- - d_1^*} + \frac{\eta}{4} \mid T_0^{(1,*)} = \infty\right) \\ & \leq \mathbb{P}\left(N_1^* \left(\frac{(\delta + \eta/4)}{b_1^- - d_1^*}\right) \log(\lceil \varepsilon K \rceil) < (\varepsilon K)^\delta \mid T_0^{(1,*)} = \infty\right) \end{aligned}$$

To estimate the above we use bounds on the transition probabilities of the conditioned process. For  $t \geq 0$  and  $i \geq 1$ , let  $\hat{P}_i(0, t)$  be the probability that there are  $i$  individuals alive at time  $t$  on the event of non-extinction. Then by Lemma 4.6.2 (see

Auxiliary results), we have

$$\begin{aligned}\hat{P}_i(0, t) &\leq \frac{\left(\int_0^t e^{-\rho(u,t)} b(u) du\right)^{i-1} \left(1 + \int_0^\infty e^{\rho(u)} d(u) ds\right)}{\left(1 + \int_0^t e^{\rho(u)} d(u) du\right) \left(1 + \int_0^t e^{-\rho(u,t)} b(u) du\right)^i} \\ &= \frac{\left(1 + \int_0^\infty e^{\rho(u)} d(u) ds\right)}{\left(1 + \int_0^t e^{\rho(u)} d(u) du\right) \left(1 + \int_0^t e^{-\rho(u,t)} b(u) du\right)} \left(1 - \frac{1}{1 + \int_0^t e^{-\rho(u,t)} b(u) du}\right)^{i-1}.\end{aligned}$$

Hence, for any  $j \geq 1$ ,

$$\begin{aligned}\mathbb{P}(N_1^*(t) < j \mid T_0^{(1,*)} = \infty) &= \sum_{i=1}^{j-1} \hat{P}_i(0, t) \\ &\leq \frac{\left(1 + \int_0^\infty e^{\rho(u)} d(u) ds\right)}{\left(1 + \int_0^t e^{\rho(u)} d(u) du\right) \left(1 + \int_0^t e^{-\rho(u,t)} b(u) du\right)} \sum_{i=1}^{j-1} \left(1 - \frac{1}{1 + \int_0^t e^{-\rho(u,t)} b(u) du}\right)^{i-1} \\ &= \frac{\left(1 + \int_0^\infty e^{\rho(u)} d(u) ds\right) \left(1 + \int_0^t e^{-\rho(u,t)} b(u) du\right)}{\left(1 + \int_0^t e^{\rho(u)} d(u) du\right) \left(1 + \int_0^t e^{-\rho(u,t)} b(u) du\right)} \left(1 - \left(\frac{\int_0^t e^{-\rho(u,t)} b(u) du}{1 + \int_0^t e^{-\rho(u,t)} b(u) du}\right)^{j-1}\right) \\ &= \frac{\left(1 + \int_0^\infty e^{\rho(u)} d(u) ds\right)}{\left(1 + \int_0^t e^{\rho(u)} d(u) du\right)} \left(1 - \left(\frac{\int_0^t e^{-\rho(u,t)} b(u) du}{1 + \int_0^t e^{-\rho(u,t)} b(u) du}\right)^{j-1}\right) \\ &\leq R_2 \left(1 - \left(\frac{\int_0^t e^{-\rho(u,t)} b(u) du}{1 + \int_0^t e^{-\rho(u,t)} b(u) du}\right)^{j-1}\right)\end{aligned}$$

for some finite  $R_2 > 0$  due to the fact that in our case  $\int_0^\infty e^{\rho(u)} d(u) ds \leq d_*/(b_- - d_*)$  and for any  $t \geq 0$ ,  $\int_0^t e^{\rho(u)} d(u) du \geq 0$ . Take  $j = \lceil (\varepsilon K)^\delta \rceil$  and note that for any  $\nu > 0$ , choosing  $x$  sufficiently close to 0, we have  $\log(1 - x) \geq -(1 + \nu)x$ . Also take



$$t = \left( \frac{\delta + \eta/4}{b_1^- - d_1^*} \right) \log(\lceil \varepsilon K \rceil)$$

$$\begin{aligned} \left( \frac{\int_0^t e^{-\rho(u,t)} b(u) du}{1 + \int_0^t e^{-\rho(u,t)} b(u) du} \right)^{j-1} &= \left( 1 - \frac{1}{1 + \int_0^t e^{-\rho(u,t)} b(u) du} \right)^{j-1} \\ &= \exp \left\{ (j-1) \log \left( 1 - \frac{1}{1 + \int_0^t e^{-\rho(u,t)} b(u) du} \right) \right\} \\ &\geq \exp \left\{ - (1 + \nu) \frac{j-1}{1 + \int_0^t e^{-\rho(u,t)} b(u) du} \right\} \end{aligned}$$

Now note that

$$\rho(u, t) = \int_u^t (d(s) - b(s)) ds \leq (d_1^* - b_1^-)(t - u).$$

and hence for any  $\nu > 0$ , for sufficiently large  $K$ ,

$$\begin{aligned} \int_0^t e^{-\rho(u,t)} b(u) du &\geq b_1^- \int_0^t e^{-\rho(u,t)} du \geq b_1^- e^{(b_1^- - d_1^*)t} \int_0^t e^{-u(b_1^- - d_1^*)} du \\ &= \left( \frac{b_1^-}{b_1^- - d_1^*} \right) e^{(b_1^- - d_1^*)t} (1 - e^{-t(b_1^- - d_1^*)}) \\ &= \left( \frac{b_1^-}{b_1^- - d_1^*} \right) (1 - \nu) (\lceil \varepsilon K \rceil)^{\delta + \frac{\eta}{4}} \end{aligned}$$

which means

$$\frac{j-1}{1 + \int_0^t e^{-\rho(u,t)} b(u) du} \leq \frac{\lceil \varepsilon K \rceil^\delta - 1}{1 + \left( \frac{b_1^-}{b_1^- - d_1^*} \right) (1 - \nu) (\lceil \varepsilon K \rceil)^{\delta + \frac{\eta}{4}}} \rightarrow 0$$

as  $K \rightarrow \infty$ . Therefore  $\mathbb{P}\left(N_1^* \left( \left( \frac{\delta + \eta/4}{b_1^- - d_1^*} \right) \log(\lceil \varepsilon K \rceil) \right) < (\varepsilon K)^\delta \mid T_0^{(1,*)} = \infty\right) \rightarrow 0$  and thus

$$\mathbb{P}\left( \frac{T_{(\varepsilon K)^\delta}^{(1,*)}}{\log(\lceil \varepsilon K \rceil)} < \frac{\delta}{\frac{1}{2}(b_1^- + b_1^+) - d_1^*} - \frac{\eta}{2} \mid T_0^{(1,*)} = \infty \right) \rightarrow 0$$

as  $K \rightarrow \infty$ . It follows that  $\mathbb{P}(E_K^c \mid T_0^{(1,*)} = \infty) \rightarrow 0$  as  $K \rightarrow \infty$ .

Now we seek the limit of  $\mathbb{P}\left( \left| \frac{T_{\varepsilon K, \delta}^{(1,*)}}{\log(\lceil \varepsilon K \rceil)} - \frac{1-\delta}{\frac{1}{2}(b_1^- + b_1^+) - d_1^*} \right| > \frac{\eta}{2} \mid T_0^{(1,*)} = \infty, E_K \right)$  as  $K \rightarrow \infty$ .

Note that by the above we know that  $\mathbb{P}(E_K \mid T_0^{(1,*)} = \infty) \rightarrow 1$  and also  $\mathbb{P}(T_0^{(1,*)} = \infty) \geq \mathbb{P}(\tau_0^{(1,*, -)} = \infty) \rightarrow 1 - (d_*/b_-) > 0$  as  $K \rightarrow \infty$ . Hence for some  $R_3 > 0$ , we have

$$\mathbb{P}\left( \left| \frac{T_{\varepsilon K, \delta}^{(1,*)}}{\log(\lceil \varepsilon K \rceil)} - \frac{1-\delta}{\frac{1}{2}(b_1^- + b_1^+) - d_1^*} \right| > \frac{\eta}{2} \mid T_0^{(1,*)} = \infty, E_K \right) \leq R_3 \mathbb{P}\left( \left| \frac{T_{\varepsilon K, \delta}^{(1,*)}}{\log(\lceil \varepsilon K \rceil)} - \frac{1-\delta}{\frac{1}{2}(b_1^- + b_1^+) - d_1^*} \right| > \right)$$

$\frac{\eta}{2} \mid E_K$ ). So on the event that the population starts with  $K^\delta$  individuals, we need to show that  $T_{\varepsilon K, \delta}^{(1,*)} / \log(\lceil \varepsilon K \rceil)$  converges in probability to  $(1-\delta) / (\frac{1}{2}(b_1^- + b_1^+) - d_1^*)$ . Given  $\delta > 0$  as above, pick  $T > \frac{1-\delta}{b_1^- - d_1^*}$ . Then for any sequence of times  $(t_K)_K$  such that for some  $c > 0$ ,  $\mathbb{P}(c \log K \leq t_K \leq T \log K) \rightarrow 1$  as  $K \rightarrow \infty$ , applying Lemma (4.2.1) with our choice of  $T$  and the continuous mapping theorem shows that

$$\frac{\log(N_1^{1,*}(t_K))}{t_K} + \frac{\rho_1^*(t_K)}{t_K} - \frac{\delta \log(\varepsilon K)}{t_K} \rightarrow 0$$

in probability  $K \rightarrow \infty$ . Consider the quantity

$$\frac{\rho_1^*(t_K)}{t_K} = \frac{1}{t_K} \int_0^{t_K} (d_1^*(s) - b_1(s)) ds.$$

Assuming  $S_K \rightarrow \infty$  as  $K \rightarrow \infty$ , for each  $K$  one can write  $t_K = 2m_K S_K + R_K S_K$  where  $(m_K)_K$  is a sequence of nonnegative integers and  $R_K \in [0, 2)$ . In the case of constant death rate, that is  $d_1^*(t) = d_1^*$ , the above quantity becomes

$$\begin{aligned} \frac{\rho_1^*(t_K)}{t_K} &= \frac{1}{(2m_K + R_K)S_K} \int_0^{2m_K S_K + R_K S_K} (d_1^* - b_1(s)) ds \\ &= \frac{m_K S_K (d_1^* - b_1^+) + m_K S_K (d_1^* - b_1^-)}{(2m_K + R_K)S_K} + \frac{1}{(2m_K + R_K)S_K} \int_0^{R_K S_K} (d_1^* - b(s)) ds \\ &= \left( d_1^* - \frac{b_1^+ + b_1^-}{2} \right) \left( 1 - \frac{R_K}{2m_K} \right) + \frac{1}{(2m_K + R_K)S_K} \int_0^{R_K S_K} (d_1^* - b(s)) ds \end{aligned}$$

Note that since  $R_K \in [0, 2)$ , the integral  $\int_0^{R_K S_K} (d_1^* - b(s)) ds$  is bounded below by 0 and bounded above by  $R_K S_K (d_1^* - b_1^-) \leq 2S_K (d_1^* - b_1^-)$ . Thus as long as  $m_K \rightarrow \infty$  as  $K \rightarrow \infty$  we get  $\frac{1}{t_K} \rho_1^*(t_K) \rightarrow d_1^* - \frac{1}{2}(b_1^+ + b_1^-)$ . This is equivalent to requiring  $\frac{t_K}{S_K} \rightarrow \infty$  as  $K \rightarrow \infty$ . Now specify that  $t_K = T_{\varepsilon K}^{(1,*)} = \inf\{t \geq 0 : N_1^*(t) \geq \varepsilon K\}$  and define for  $\star \in \{-, +\}$ , the process  $N_{1,\delta}^{*,\star}(t)$  as

$$N_{1,\delta}^{*,\star}(t) = (\varepsilon K)^\delta + \int_0^\infty \int_0^t \left( \mathbb{1}_{\{0 \leq \theta \leq b_1^* N_{1,\delta}^{*,\star}(s)\}} - \mathbb{1}_{\{b_1^* N_{1,\delta}^{*,\star}(s) \leq \theta \leq (b_1^* + d_1^*) N_{1,\delta}^{*,\star}(s)\}} \right) Q_1(ds, d\theta),$$

using the same Poisson random measure  $Q_1(ds, d\theta)$  as the process  $N_{1,\delta}^*(t)$ . Similarly to the first part,  $N_{1,\delta}^{*,\star}(t)$  is a standard birth-death process with fixed birth rate  $b_1^*$  and fixed death rate  $d_1^*$  and since  $b_1^- \leq b_1(t) \leq b_1^+$ , if  $\tau_{\varepsilon K, \delta}^{1,*,\star} = \inf\{t \geq 0 : N_{1,\delta}^{*,\star}(t) \geq \varepsilon K\}$  we have that  $\tau_{\varepsilon K, \delta}^{1,*,+} < T_{\varepsilon K, \delta}^{(1,*)}$  almost surely. By Corollary 4.1.3 we have

$$\frac{\tau_{\varepsilon K, \delta}^{1,*,+}}{\log(\lceil \varepsilon K \rceil)} \rightarrow \frac{1-\delta}{b_1^+ - d_1^*}$$

as  $K \rightarrow \infty$ , in probability. This implies there exists a constant  $C > 0$  such that  $\mathbb{P}(C \log K \leq T_{\varepsilon K, \delta}^{(1,*)}) \rightarrow 1$  as  $K \rightarrow \infty$ . So for this choice of  $(t_K)_K$ , on the event  $C \log K \leq T_{\varepsilon K, \delta}^{(1,*)}$ , the condition  $\frac{t_K}{S_K} \rightarrow \infty$  is satisfied if

$$\frac{S_K}{\log K} \rightarrow 0$$

as  $K \rightarrow \infty$ . By our choice of  $(t_K)_K$  we have

$$\frac{\log(N_1^*(t_K))}{t_K} = \frac{\log(\lceil \varepsilon K \rceil)}{T_{\varepsilon K, \delta}^{(1,*)}}$$

which means, assuming  $S_K / \log K \rightarrow 0$  as  $K \rightarrow \infty$ ,

$$\frac{\log(\lceil \varepsilon K \rceil)}{T_{\varepsilon K, \delta}^{(1,*)}} (1 - \delta) \rightarrow \frac{1}{2}(b_1^- + b_1^+) - d_1^*$$

in probability, which implies

$$\frac{T_{\varepsilon K, \delta}^{(1,*)}}{\log(\lceil \varepsilon K \rceil)} \rightarrow \frac{1 - \delta}{\frac{1}{2}(b_1^- + b_1^+) - d_1^*}$$

in probability. □

### 4.3 First Phase

Before proving the main theorem of this section, we need the following lemma on the probability of extinction of the periodic birth-death processes  $N_1^-(t)$  and  $N_1^+(t)$ . Note that in this section we do not assume the adjusted death rate functions  $d_1^*(t)$  are necessarily constant.

**Lemma 4.3.1.** *Consider either of the periodic birth-death processes  $N_1^*(t)$  for  $* \in \{-, +\}$ . Suppose in all cases that  $S_K / \log K \rightarrow 0$  and for all  $* \in \{-, +\}$  assume that  $\sup_{K \geq 1} \sup_{t \geq 0} (d_1^*(t) - b_1(t)) < 0$ .*

(i) *For any sequence  $(t_K)_K$  with  $t_K \rightarrow \infty$  as  $K \rightarrow \infty$ ,*

$$\mathbb{P}(t_K < T_0^{(1,*)} < \infty) = \int_{t_K}^{\infty} e^{\rho_1^*(s)} d_1^*(s) ds \rightarrow 0$$

(ii) *Suppose  $S_K \rightarrow \infty$  as  $K \rightarrow \infty$ . If additionally for all  $* \in \{-, +\}$  and  $t \geq 0$  the*

limit  $\lim_{K \rightarrow \infty} \rho_1^*(t)$  exists pointwise, then as  $K \rightarrow \infty$ ,

$$\mathbb{P}(T_0^{(1,*)} < \infty) = \frac{\int_0^\infty e^{\rho_1^*(s)} d_1^*(s) ds}{1 + \int_0^\infty e^{\rho_1^*(s)} d_1^*(s) ds} \rightarrow \frac{d_1^{*,+}}{b_1^+}$$

(iii) Suppose  $S_K \rightarrow 0$  as  $K \rightarrow \infty$ . Then as  $K \rightarrow \infty$ ,

$$\mathbb{P}(T_0^{(1,*)} < \infty) = \frac{\int_0^\infty e^{\rho_1^*(s)} d_1^*(s) ds}{1 + \int_0^\infty e^{\rho_1^*(s)} d_1^*(s) ds} \rightarrow \frac{d_1^{*,+} + d_1^{*, -}}{b_1^+ + b_1^-}$$

*Proof.* (i) Equation (18) of [24] shows that for a time-varying birth-death process with birth intensity  $b(t)$ , death intensity  $d(t)$  and extinction time  $T_0$ , one has

$$\mathbb{P}(T_0 \leq t) = \frac{\int_0^t e^{\rho(s)} d(s) ds}{1 + \int_0^t e^{\rho(s)} d(s) ds}$$

where  $\rho(s) = \int_0^s (d(u) - b(u)) du$ . Assuming  $\sup_{K \geq 1} \sup_{t \geq 0} (d_1^*(t) - b_1(t)) < 0$  one can choose a positive constant  $C_* < -\sup_{K \geq 1} \sup_{t \geq 0} (d_1^*(t) - b_1(t))$ . Additionally since we assume  $d_1^*(t)$  is bounded (uniformly in  $K$ ), we see that  $e^{\rho_1^*(s)} d_1^*(s) \leq d_{\max}^* e^{-C_* s}$ . Hence,

$$\begin{aligned} \mathbb{P}(t_K < T_0^{(1,*)} < \infty) &= \frac{\int_0^\infty e^{\rho_1^*(s)} d_1^*(s) ds}{1 + \int_0^\infty e^{\rho_1^*(s)} d_1^*(s) ds} - \frac{\int_0^{t_K} e^{\rho_1^*(s)} d_1^*(s) ds}{1 + \int_0^{t_K} e^{\rho_1^*(s)} d_1^*(s) ds} \\ &\leq \frac{\int_0^\infty e^{\rho_1^*(s)} d_1^+(s) ds}{1 + \int_0^\infty e^{\rho_1^*(s)} d_1^*(s) ds} - \frac{\int_0^{t_K} e^{\rho_1^*(s)} d_1^*(s) ds}{1 + \int_0^\infty e^{\rho_1^*(s)} d_1^*(s) ds} \\ &\leq \int_{t_K}^\infty e^{\rho_1^*(s)} d_1^*(s) ds \leq \int_{t_K}^\infty d_{\max}^* e^{-C_* s} ds \\ &= \frac{d_{\max}^*}{C_*} \left( e^{-C_* t_K} \right) \rightarrow 0 \text{ as } K \rightarrow \infty. \end{aligned}$$

(ii) To calculate  $\mathbb{P}(T_0^{(1,*)} < \infty)$  for  $* \in \{-, +\}$ , we seek the value of the limit  $\lim_{K \rightarrow \infty} \int_0^\infty e^{\rho_1^*(s)} d_1^*(s) ds$ . As before, assuming that  $\sup_{K \geq 1} \sup_{t \geq 0} (d_1^*(t) - b_1(t)) < 0$  one can choose a positive constant  $C_* < -\sup_{K \geq 1} \sup_{t \geq 0} (d_1^*(t) - b_1(t))$  which means that  $d_{\max}^* e^{-C_* s}$  is an inte-

grable upper bound on the nonnegative function  $e^{\rho_1^*(s)} d_1^*(s)$ . If the pointwise limit as  $K \rightarrow \infty$  of this function exists, by the Dominated Convergence Theorem we may pass the limit inside the integral, so in the case that for each  $* \in \{-, +\}$ ,  $d_1^*(t) \leq d_{\max}^*$ , we only need to consider the limit  $\lim_{K \rightarrow \infty} \rho_1^*(s)$ . In the particular case that the period length goes to infinity as  $K \rightarrow \infty$  and the pair  $(b_1(t), d_1^*(t))$  takes constant values  $(b_1^+, d_1^{*,+})$  and  $(b_1^-, d_1^{*, -})$  in odd and even periods respectively, then  $\lim_{K \rightarrow \infty} \rho_1^*(s) = s(d_1^{*,+} - b_1^+) < 0$  and hence  $\lim_{K \rightarrow \infty} \int_0^\infty e^{\rho_1^*(s)} d_1^*(s) ds = d_1^{*,+} / (b_1^+ - d_1^{*,+})$ . Therefore

$$\mathbb{P}(T_0^{(1,*)} < \infty) = \frac{\int_0^\infty e^{\rho_1^*(s)} d_1^*(s) ds}{1 + \int_0^\infty e^{\rho_1^*(s)} d_1^*(s) ds} \rightarrow \frac{d_1^{*,+}}{b_1^+}$$

as  $K \rightarrow \infty$ .

(iii) As in part (ii), we seek the value of the limit  $\lim_{K \rightarrow \infty} \int_0^\infty e^{\rho_1^*(s)} d_1^*(s) ds$ . We will first prove a more general result. Suppose that  $\sup_{K \geq 1} \sup_{t \geq 0} d_K(t) \leq d_{\max}$  and  $\sup_{K \geq 1} \sup_{t \geq 0} b_K(t) \leq b_{\max}$  for some  $b_{\max}, d_{\max} > 0$ . Assume also that for some  $\varepsilon > 0$ ,

$$\sup_{K \geq 1} \sup_{t \geq 0} \frac{\rho_K(t)}{t} \leq -\varepsilon. \quad (4.5)$$

Define the averaged quantities  $\bar{b} = \frac{1}{2S_K} \int_0^{2S_K} b_K(s) ds$ ,  $\bar{d} = \frac{1}{2S_K} \int_0^{2S_K} d_K(s) ds$  and  $\bar{\rho} = \bar{d} - \bar{b}$ . Then choose a sequence  $(m_K)_K$  of positive integers such that  $m_K S_K \rightarrow \infty$  as  $K \rightarrow \infty$ . We have

$$\begin{aligned} \left| \int_0^\infty e^{\rho_K(s)} d_K(s) ds - \int_0^\infty e^{\bar{\rho}s} \bar{d} ds \right| &\leq \left| \int_0^{2m_K S_K} e^{\rho_K(s)} d_K(s) ds - \int_0^{2m_K S_K} e^{\bar{\rho}s} \bar{d} ds \right| \\ &\quad + \left| \int_{2m_K S_K}^\infty e^{\rho_K(s)} d_K(s) ds + \int_{2m_K S_K}^\infty e^{\bar{\rho}s} \bar{d} ds \right| \end{aligned} \quad (4.6)$$

For the second term in (4.6), using (4.5) we have

$$\left| \int_{2m_K S_K}^\infty e^{\rho_K(s)} d_K(s) ds + \int_{2m_K S_K}^\infty e^{\bar{\rho}s} \bar{d} ds \right| \leq 2d_{\max} \int_{2m_K S_K}^\infty e^{-\varepsilon s} ds \rightarrow 0$$

as  $K \rightarrow \infty$ . We bound the first term in (4.6) as follows:

$$\begin{aligned} \left| \int_0^{2m_K S_K} e^{\rho_K(s)} d_K(s) - \int_0^{2m_K S_K} e^{\bar{\rho}s} \bar{d} ds \right| &\leq \left| \int_0^{2m_K S_K} e^{\rho_K(s)} d_K(s) - \int_0^{2m_K S_K} e^{\bar{\rho}s} d_K(s) ds \right| \\ &\quad + \left| \int_0^{2m_K S_K} e^{\bar{\rho}s} (d_K(s) - \bar{d}) ds \right| \end{aligned} \quad (4.7)$$

For the first term of (4.7), note that  $|\rho_K(s) - \bar{\rho}s| \leq 2S_K(d_{\max} + b_{\max})$ . Therefore,

$$|e^{\rho_K(s)} - e^{\bar{\rho}s}| \leq e^{\bar{\rho}s} |e^{\rho_K(s) - \bar{\rho}s} - 1| \leq e^{\bar{\rho}s} (e^{2S_K(d_{\max} + b_{\max})} - 1) \leq CS_K$$

for sufficiently small  $S_K$  and some  $C > 0$ . Hence

$$\left| \int_0^{2m_K S_K} e^{\rho_K(s)} d_K(s) - \int_0^{2m_K S_K} e^{\bar{\rho}s} d_K(s) ds \right| \leq CS_K d_{\max} \int_0^{2m_K S_K} e^{\bar{\rho}s} ds \rightarrow 0$$

as  $K \rightarrow \infty$ . Now we turn to the second term in (4.7). Observe that for integers  $0 \leq i \leq m_K$ ,

$$\begin{aligned} &\left| \int_{2iS_K}^{2(i+1)S_K} e^{\bar{\rho}s} (d_K(s) - \bar{d}) ds \right| \\ &\leq \left| \int_{2iS_K}^{2(i+1)S_K} (e^{\bar{\rho}s} - e^{2iS_K \bar{\rho}}) (d_K(s) - \bar{d}) ds \right| + e^{2iS_K \bar{\rho}} \left| \int_{2iS_K}^{2(i+1)S_K} (d_K(s) - \bar{d}) ds \right| \\ &\leq d_{\max} \int_{2iS_K}^{2(i+1)S_K} e^{2iS_K \bar{\rho}} (e^{\bar{\rho}s - 2iS_K \bar{\rho}} - 1) ds \\ &\leq 2d_{\max} S_K e^{2iS_K \bar{\rho}} (e^{2S_K \bar{\rho}} - 1) \end{aligned}$$

Hence the second term in (4.7) satisfies

$$\begin{aligned} \left| \int_0^{2m_K S_K} e^{\bar{\rho}s} (d_K(s) - \bar{d}) ds \right| &\leq \sum_{i=0}^{m_K-1} \left| \int_{2iS_K}^{2(i+1)S_K} e^{\bar{\rho}s} (d_K(s) - \bar{d}) ds \right| \\ &\leq 2d_{\max} S_K (e^{2S_K \bar{\rho}} - 1) \sum_{i=0}^{m_K-1} e^{2iS_K \bar{\rho}} \\ &\leq 2d_{\max} S_K (e^{2S_K \bar{\rho}} - 1) \left( \frac{1}{1 - e^{-2S_K \bar{\rho}}} \right) \\ &\leq 2d_{\max} S_K \rightarrow 0 \end{aligned}$$

as  $K \rightarrow \infty$ . In the specific case we consider, the condition (4.5) is automatically satisfied and the averaged quantities become  $\bar{b} = \frac{1}{2}(b_1^+ + b_1^-)$ ,  $\bar{d} = \frac{1}{2}(d_1^{*, -} + d_1^{*, +})$  and

$$\bar{\rho} = \frac{1}{2}((d_1^{*,+} - b_1^+) + (d_1^{*, -} - b_1^-)). \quad \square$$

Note that parts (ii) and (iii) of the above lemma also hold for the limit  $\lim_{K \rightarrow \infty} \mathbb{P}(T_0 < t_K)$  provided  $t_K \rightarrow \infty$  as  $K \rightarrow \infty$ . The next result (Theorem 1.2.1 from the introduction restated) covers the first phase of invasion and determines both the probability of invasion and the duration of the phase.

**Theorem 4.3.2.** *Let the period  $S_K$  be such that  $S_K/\log K \rightarrow 0$  as  $K \rightarrow \infty$ . Suppose also that  $d_1(t) = d_1 < b_1^-$ , a constant, for all  $t \geq 0$ .*

(a) *If  $S_K \rightarrow \infty$  as  $K \rightarrow \infty$ , there exist constants  $R$  and  $R'$  such that for any  $\varepsilon > 0$ ,*

$$(i) \quad \mathbb{P}(T_0^{(1)} < \tilde{T}_\varepsilon^{(0)}) \rightarrow \frac{d_1}{b_1^+} - R\varepsilon \quad \text{as } K \rightarrow \infty$$

$$(ii) \quad \mathbb{P}(T_{\varepsilon K}^{(1)} < \tilde{T}_\varepsilon^{(0)}) \rightarrow \left(1 - \frac{d_1^-}{b_1^+}\right) - R\varepsilon = \left(1 - \frac{d_1}{b_1^+}\right) - R'\varepsilon \quad \text{as } K \rightarrow \infty$$

(b) *If instead  $S_K \rightarrow 0$  as  $K \rightarrow \infty$ , there exist constants  $R$  and  $R'$  such that for any  $\varepsilon > 0$ ,*

$$(i) \quad \mathbb{P}(T_0^{(1)} < \tilde{T}_\varepsilon^{(0)}) \rightarrow \frac{d_1}{\frac{1}{2}(b_1^+ + b_1^-)} - R\varepsilon \quad \text{as } K \rightarrow \infty$$

$$(ii) \quad \mathbb{P}(T_{\varepsilon K}^{(1)} < \tilde{T}_\varepsilon^{(0)}) \rightarrow \left(1 - \frac{d_1^-}{\frac{1}{2}(b_1^+ + b_1^-)}\right) - R\varepsilon = \left(1 - \frac{d_1}{\frac{1}{2}(b_1^+ + b_1^-)}\right) - R'\varepsilon \quad \text{as } K \rightarrow \infty$$

(c) *Moreover, there exists  $\beta > 0$  such that for sufficiently large  $K$ ,*

$$\mathbb{P}\left(\left(1 - c\varepsilon\right)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1} < T_{\varepsilon K}^{(1)} < (1 + c\varepsilon)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1} \mid T_{\varepsilon K}^{(1)} < \tilde{T}_\varepsilon^{(0)}\right) \geq 1 - \beta\varepsilon.$$

*Proof.* We prove parts (a) and (b) simultaneously. From the way the approximating processes are defined, we have almost surely:

$$N_1^-(t) \leq N_1(t) \leq N_1^+(t) \quad \text{for all } t \in [0, \tilde{T}_\varepsilon^{(0)}]$$

and

$$N_0^-(t) \leq N_0(t) \leq N_0^+(t) \quad \text{for all } t \in [0, T_{\varepsilon K}^{(1)}].$$

Define the event

$$A_\varepsilon^K = \{T_{\varepsilon K}^{(1)} \wedge T_0^{(1)} < \tilde{T}_\varepsilon^{(0)}\}.$$

Hence for part (i),

$$\begin{aligned}
\mathbb{P}(T_0^{(1)} < \tilde{T}_\varepsilon^{(0)}) &\geq \mathbb{P}(T_0^{(1)} < \tilde{T}_\varepsilon^{(0)}, A_\varepsilon^K) \\
&\geq \mathbb{P}(T_0^{(1)} < T_{\varepsilon K}^{(1)}, A_\varepsilon^K) \\
&\geq \mathbb{P}(T_0^{(1,+)} < T_{\varepsilon K}^{(1,+)}, A_\varepsilon^K) \\
&\geq \mathbb{P}(T_0^{(1,+)} < T_{\varepsilon K}^{(1,+)} - \mathbb{P}((A_\varepsilon^K)^c).
\end{aligned} \tag{4.8}$$

**Part (i), Step 1: Bounding  $\mathbb{P}((A_\varepsilon^K)^c)$**

The first aim is to find an upper bound for  $\mathbb{P}((A_\varepsilon^K)^c)$ . Afterwards we can then assume that the type 0 population stays close to its equilibrium (i.e. within the set  $I_\varepsilon^{(0)}$ ) with high probability. By Lemma 4.1.4, for any  $\varepsilon > 0$  and each  $* \in \{-, +\}$  there exists a constant  $V_* > 0$  such that

$$\mathbb{P}(\tilde{T}_\varepsilon^{(0,*)} > e^{V_*K}) \rightarrow 1 \text{ as } K \rightarrow \infty. \tag{4.9}$$

Choose  $V = V_- \wedge V_+$  and write

$$\mathbb{P}((A_\varepsilon^K)^c) = \mathbb{P}(\tilde{T}_\varepsilon^{(0)} < T_{\varepsilon K}^{(1)} \wedge T_0^{(1)}, \tilde{T}_\varepsilon^{(0)} \leq e^{VK}) + \mathbb{P}(e^{VK} < \tilde{T}_\varepsilon^{(0)} < T_{\varepsilon K}^{(1)} \wedge T_0^{(1)}).$$

For the first term, note that

$$\begin{aligned}
\mathbb{P}(\tilde{T}_\varepsilon^{(0)} < T_{\varepsilon K}^{(1)} \wedge T_0^{(1)}, \tilde{T}_\varepsilon^{(0)} \leq e^{VK}) &\leq \mathbb{P}(\tilde{T}_\varepsilon^{(0)} < T_{\varepsilon K}^{(1)}, \tilde{T}_\varepsilon^{(0)} \leq e^{VK}) \\
&\leq \mathbb{P}(\tilde{T}_\varepsilon^{(0,+)} \wedge \tilde{T}_\varepsilon^{(0,-)} \leq e^{VK}) \rightarrow 0 \text{ as } K \rightarrow \infty
\end{aligned}$$

due to (4.9). For the second term, note that while both the key couplings hold,  $T_0^{(1)} \leq T_0^{(1,+)}$  and  $T_{\varepsilon K}^{(1)} \leq T_{\varepsilon K}^{(1,-)}$ , almost surely. Hence,

$$\begin{aligned}
\mathbb{P}(e^{VK} < \tilde{T}_\varepsilon^{(0)} < T_{\varepsilon K}^{(1)} \wedge T_0^{(1)}) &\leq \mathbb{P}(e^{VK} < T_{\varepsilon K}^{(1,-)} \wedge T_0^{(1,+)}, \tilde{T}_\varepsilon^{(0)} > e^{VK}) \\
&\leq \mathbb{P}(e^{VK} < T_{\varepsilon K}^{(1,-)} \wedge T_0^{(1,+)})) \\
&= \mathbb{P}(e^{VK} < T_{\varepsilon K}^{(1,-)}, e^{VK} < T_0^{(1,+)}, T_0^{(1,-)} < \infty, T_0^{(1,+)} < \infty) \\
&+ \mathbb{P}(e^{VK} < T_{\varepsilon K}^{(1,-)}, e^{VK} < T_0^{(1,+)}, T_0^{(1,-)} = \infty, T_0^{(1,+)} < \infty) \\
&+ \mathbb{P}(e^{VK} < T_{\varepsilon K}^{(1,-)}, e^{VK} < T_0^{(1,+)}, T_0^{(1,-)} < \infty, T_0^{(1,+)} = \infty) \\
&+ \mathbb{P}(e^{VK} < T_{\varepsilon K}^{(1,-)}, e^{VK} < T_0^{(1,+)}, T_0^{(1,-)} = \infty, T_0^{(1,+)} = \infty).
\end{aligned} \tag{4.10}$$

The aim is now to bound each of the above four probabilities in terms of  $K$ . Consider



the first of the four probabilities. We have

$$\begin{aligned} & \mathbb{P}(e^{VK} < T_{\varepsilon K}^{(1,-)}, e^{VK} < T_0^{(1,+)}, T_0^{(1,-)} < \infty, T_0^{(1,+)} < \infty) \\ & \leq \mathbb{P}(e^{VK} < T_0^{(1,+)}, T_0^{(1,+)} < \infty) = \mathbb{P}(T_0^{(1,+)} < \infty) - \mathbb{P}(T_0^{(1,+)} \leq e^{VK}). \end{aligned}$$

Equation (18) of [24] shows that for a time-varying birth-death process with birth intensity  $b(t)$  and death intensity  $d(t)$  one has

$$\mathbb{P}(T_0 \leq t) = \frac{\int_0^t e^{\rho(s)} d(s) ds}{1 + \int_0^t e^{\rho(s)} d(s) ds}$$

where  $\rho(s) = \int_0^s (d(u) - b(u)) du$ . In our case,

$$\sup_{K \geq 1} \sup_{t \geq 0} (d_1^+(t) - b_1(t)) = d_1 - b_1^- < 0,$$

so defining  $t_K = e^{VK}$  we have

$$\mathbb{P}(T_0^{(1,+)} < \infty) - \mathbb{P}(T_0^{(1,+)} \leq e^{VK}) = \mathbb{P}(t_K < T_0^{(1,+)} < \infty) \rightarrow 0 \text{ as } K \rightarrow \infty.$$

using part (i) of Lemma 4.3.1. Next observe that the second probability in (2) is zero for any fixed  $K$  since by definition of the type 1 coupling, the event  $\{T_0^{(1,-)} = \infty, T_0^{(1,+)} < \infty\}$  has probability zero. For the third probability note that

$$\begin{aligned} & \mathbb{P}(e^{VK} < T_{\varepsilon K}^{(1,-)}, e^{VK} < T_0^{(1,+)}, T_0^{(1,-)} < \infty, T_0^{(1,+)} = \infty) \\ & \leq \mathbb{P}(T_0^{(1,-)} < \infty, T_0^{(1,+)} = \infty) = \mathbb{P}(T_0^{(1,-)} < \infty) - \mathbb{P}(T_0^{(1,+)} < \infty). \end{aligned}$$

From Lemma 4.3.1 parts (ii) and (iii) we have

$$\mathbb{P}(T_0^{(1,-)} < \infty) - \mathbb{P}(T_0^{(1,+)} < \infty) \rightarrow \frac{d_1^-}{b_1^+} - \frac{d_1^+}{b_1^+} = \left( \frac{2C_{1,0}}{b_1^+ C_{0,0}} \right) \varepsilon \leq R\varepsilon$$

in the case that  $S_K \rightarrow \infty$  as  $K \rightarrow \infty$ . Moreover, if  $S_K \rightarrow 0$  as  $K \rightarrow \infty$ , by Lemma 4.3.1 part (iii),

$$\mathbb{P}(T_0^{(1,-)} < \infty) - \mathbb{P}(T_0^{(1,+)} < \infty) \rightarrow \frac{d_1^-}{\frac{1}{2}(b_1^+ + b_1^-)} - \frac{d_1^+}{\frac{1}{2}(b_1^+ + b_1^-)} = \left( \frac{C_{1,0}}{(b_1^+ + b_1^-)C_{0,0}} \right) \varepsilon \leq R\varepsilon$$

where the constant  $R > 0$  does not depend on  $K$ . Lastly, for the fourth of the proba-

bilities in (2) we have

$$\begin{aligned} \mathbb{P}(e^{VK} < T_{\varepsilon K}^{(1,-)}, e^{VK} < T_0^{(1,+)}, T_0^{(1,-)} = \infty, T_0^{(1,+)} = \infty) &\leq \mathbb{P}(e^{VK} < T_{\varepsilon K}^{(1,-)}, T_0^{(1,-)} = \infty) \\ &\rightarrow 0 \end{aligned}$$

as  $K \rightarrow \infty$ , since on the event  $T_0^{(1,-)} = \infty$  for each  $\eta > 0$  we can choose  $K$  large enough so that  $\mathbb{P}\left(T_{\varepsilon K}^{(1,-)} \leq \log(\lceil \varepsilon K \rceil)(b_1^- - d_1^-)^{-1} + \eta < e^{VK}\right) \rightarrow 1$  as  $K \rightarrow \infty$ . Hence,

$$\lim_{K \rightarrow \infty} \mathbb{P}((A_\varepsilon^K)^c) \leq R\varepsilon. \quad (4.11)$$

### Part (i), Step 2: Using the Time to Reach a Level

In view of (4.8) and (4.11), for large enough  $K$ , given a sequence  $(t_K)_K$  of increasing positive real numbers such that  $t_K \leq a \log(\varepsilon K)$  for some  $a < \frac{1}{b_1^+ - d_1^+}$  and  $t_K \rightarrow \infty$  as  $K \rightarrow \infty$ , we have

$$\begin{aligned} \mathbb{P}(T_0^{(1)} < \tilde{T}_\varepsilon^{(0)}) &\geq \mathbb{P}(T_0^{(1,+)} < T_{\varepsilon K}^{(1,+)} - R\varepsilon) \\ &\geq \mathbb{P}(T_0^{(1)} < t_K < T_{\varepsilon K}^{(1,+)} - R\varepsilon) \\ &= \mathbb{P}(T_0^{(1,+)} < t_K) - \mathbb{P}(T_{\varepsilon K}^{(1,+)} < t_K) - R\varepsilon \end{aligned}$$

By using the fact  $T_{\varepsilon K}^{(1,+)} < T_{\varepsilon K}^{(1,+,+)}$  almost surely, where  $T_{\varepsilon K}^{(1,+,+)}$  is the time taken to reach level  $\varepsilon K$  by the time-homogeneous birth-death process with birth rate  $b_1^+$  and death rate  $d_1^+$ , we get that

$$\mathbb{P}(T_{\varepsilon K}^{(1,+)} < t_K) \leq \mathbb{P}(T_{\varepsilon K}^{(1,+,+)} < t_K) \rightarrow 0$$

as  $K \rightarrow \infty$  (due to the fact that by part 2 of Lemma 4.1.1, on the event of non-extinction  $T_{\varepsilon K}^{(1,+,+)}/\log(\varepsilon K) \rightarrow 1/(b_1^+ - d_1^+)$  in probability as  $K \rightarrow \infty$ ). Thus using parts (ii) and (iii) of Lemma 4.3.1 we get the limit as:

$$\mathbb{P}(T_0^{(1)} < \tilde{T}_\varepsilon^{(0)}) \rightarrow \frac{d_1^+}{b_1^+} - R\varepsilon,$$

where  $S_K \rightarrow \infty$  as  $K \rightarrow \infty$ , and

$$\mathbb{P}(T_0^{(1)} < \tilde{T}_\varepsilon^{(0)}) \rightarrow \frac{d_1^+}{\frac{1}{2}(b_1^+ + b_1^-)} - R\varepsilon,$$

where  $S_K \rightarrow 0$  as  $K \rightarrow \infty$ .

**Part (ii)**

For part (ii), note that

$$\begin{aligned}
\mathbb{P}(T_{\varepsilon K}^{(1)} < \tilde{T}_\varepsilon^{(0)}) &\geq \mathbb{P}(T_{\varepsilon K}^{(1)} < \tilde{T}_\varepsilon^{(0)}, A_\varepsilon^K) \\
&\geq \mathbb{P}(T_{\varepsilon K}^{(1)} < T_0^{(1)}, A_\varepsilon^K) \\
&\geq \mathbb{P}(T_{\varepsilon K}^{(1,-)} < T_0^{(1,-)}, A_\varepsilon^K) \\
&\geq \mathbb{P}(T_{\varepsilon K}^{(1,-)} < T_0^{(1,-)}) - \mathbb{P}((A_\varepsilon^K)^c) \\
&\geq \mathbb{P}(T_{\varepsilon K}^{(1,-)} < T_0^{(1,-)}) - R\varepsilon
\end{aligned}$$

where the same calculation as for part (i) enables the last inequality. Analogously to part (i) we obtain

$$\mathbb{P}(T_{\varepsilon K}^{(1,-)} < T_0^{(1,-)}) \rightarrow 1 - \frac{d_1^-}{b_1^+}$$

for the case  $S_K \rightarrow \infty$  and we obtain

$$\mathbb{P}(T_{\varepsilon K}^{(1,-)} < T_0^{(1,-)}) \rightarrow \left(1 - \frac{d_1^-}{\frac{1}{2}(b_1^+ + b_1^-)}\right)$$

for the case  $S_K \rightarrow 0$ .

**Part (c)**

For the last statement in the theorem, let  $c > 0$  and  $\varepsilon > 0$  and in view of Lemma 4.2.2 choose  $K$  large enough so that for each  $* \in \{-, +\}$ ,

$$\mathbb{P}\left((1-c\varepsilon)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1^+} < T_{\varepsilon K}^{(1,*)} < (1+c\varepsilon)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1^+} \mid T_0^{(1,*)} = \infty\right) > 1-\eta_K \tag{4.12}$$

for some sequence  $(\eta_K)_K$  such that  $\eta_K \rightarrow 0$  as  $K \rightarrow \infty$ . Using the fact that the

couplings on the type 1 process hold until time  $\tilde{T}_\varepsilon^{(0)}$ ,

$$\begin{aligned} & \mathbb{P}\left((1 - c\varepsilon)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1^+} < T_{\varepsilon K}^{(1)} < (1 + c\varepsilon)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1^+}, T_{\varepsilon K}^{(1)} < \tilde{T}_\varepsilon^{(0)}\right) \\ & \geq \mathbb{P}\left((1 - c\varepsilon)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1^+} < T_{\varepsilon K}^{(1,+)} \leq T_{\varepsilon K}^{(1,-)} < (1 + c\varepsilon)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1^+}\right) \\ & - \mathbb{P}\left((1 - c\varepsilon)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1^+} < T_{\varepsilon K}^{(1,+)} \leq T_{\varepsilon K}^{(1,-)} < (1 + c\varepsilon)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1^+}, \tilde{T}_\varepsilon^{(0)} < T_{\varepsilon K}^{(1)}\right) \end{aligned}$$

Recall that there exists  $V > 0$  such that  $\mathbb{P}(\tilde{T}_\varepsilon^{(0)} > e^{VK}) \rightarrow 1$  as  $K \rightarrow \infty$ . Note that for any such  $V$ ,

$$\begin{aligned} & \mathbb{P}\left((1 - c\varepsilon)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1^+} < T_{\varepsilon K}^{(1,+)} \leq T_{\varepsilon K}^{(1,-)} < (1 + c\varepsilon)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1^+}, \tilde{T}_\varepsilon^{(0)} < T_{\varepsilon K}^{(1)}\right) \\ & = \mathbb{P}\left((1 - c\varepsilon)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1^+} < T_{\varepsilon K}^{(1,+)} \leq T_{\varepsilon K}^{(1,-)} < (1 + c\varepsilon)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1^+}, \tilde{T}_\varepsilon^{(0)} < T_{\varepsilon K}^{(1)}\right. \\ & \left., \tilde{T}_\varepsilon^{(0)} < e^{VK}\right) \\ & + \mathbb{P}\left((1 - c\varepsilon)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1^+} < T_{\varepsilon K}^{(1,+)} \leq T_{\varepsilon K}^{(1,-)} < (1 + c\varepsilon)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1^+}\right. \\ & \left., e^{VK} < \tilde{T}_\varepsilon^{(0)} < T_{\varepsilon K}^{(1)}\right). \end{aligned}$$

By the choice of  $V$ , the first probability in the above sum converges to zero as  $K \rightarrow \infty$  and for large enough  $K$ , the event in the second probability implies  $T_{\varepsilon K}^{(1,-)} < e^{VK} < T_{\varepsilon K}^{(1)}$  and thus has probability zero. From earlier calculations involving the probability of extinction one sees that

$$\mathbb{P}\left((1 + c\varepsilon)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1^+} < T_0^{(1,-)} < \infty\right) \rightarrow 0$$

as  $K \rightarrow \infty$ . Additionally,

$$\mathbb{P}\left((1 - c\varepsilon)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1^+} < T_0^{(1,-)} < (1 + c\varepsilon)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1^+}\right) \rightarrow 0$$

and using (4.12) as well as the fact that  $\mathbb{P}(T_0^{(1,-)} < \infty) \rightarrow d_1^-/b_1^+$  as  $K \rightarrow \infty$ , we have

$$\begin{aligned}
& \mathbb{P}\left((1 - c\varepsilon)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1^+} < T_{\varepsilon K}^{(1,+)} \leq T_{\varepsilon K}^{(1,-)} < (1 + c\varepsilon)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1^+}, T_0^{(1,-)} = \infty\right) \\
&= \mathbb{P}\left((1 - c\varepsilon)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1^+} < T_{\varepsilon K}^{(1,+)} \leq T_{\varepsilon K}^{(1,-)} < (1 + c\varepsilon)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1^+} \mid T_0^{(1,-)} = \infty, T_0^{(1,+)} = \infty\right) \mathbb{P}\left(T_0^{(1,-)} = \infty\right) \\
&\rightarrow 1 - \frac{d_1^-}{b_1^+}
\end{aligned} \tag{4.13}$$

as  $K \rightarrow \infty$ . Using the fact that  $\mathbb{P}(A_\varepsilon^K) > 1 - R\varepsilon$  for large enough  $K$ ,

$$\begin{aligned}
\mathbb{P}(T_{\varepsilon K}^{(1)} < \tilde{T}_\varepsilon^{(0)}) &= \mathbb{P}(A_\varepsilon^K) - \mathbb{P}(T_0^{(1)} < \tilde{T}_\varepsilon^{(0)} < T_{\varepsilon K}^{(1)}) \\
&\geq \mathbb{P}(A_\varepsilon^K) - \mathbb{P}(T_0^{(1,-)} < \tilde{T}_\varepsilon^{(0)} < T_{\varepsilon K}^{(1,-)}) \\
&\geq 1 - R\varepsilon - \mathbb{P}(T_0^{(1,-)} < \infty) \rightarrow 1 - R\varepsilon - \frac{d_1^-}{b_1^+} > 0.
\end{aligned}$$

For large enough  $K$  we also have

$$\begin{aligned}
\mathbb{P}(T_0^{(1)} < \tilde{T}_\varepsilon^{(0)} < T_{\varepsilon K}^{(1)}) &= \mathbb{P}(T_0^{(1)} < \tilde{T}_\varepsilon^{(0)}, T_0^{(1)} < T_{\varepsilon K}^{(1)}) \\
&\geq \mathbb{P}(T_0^{(1,+)} < e^{VK} < \tilde{T}_\varepsilon^{(0)}, T_0^{(1,+)} < T_{\varepsilon K}^{(1,+)}, T_0^{(1,+)} < \infty) \\
&\geq \mathbb{P}(T_0^{(1,+)} < e^{VK} < \tilde{T}_\varepsilon^{(0)}, T_0^{(1,+)} < T_{\varepsilon K}^{(1,+)} \mid T_0^{(1,+)} < \infty) \mathbb{P}(T_0^{(1,+)} < \infty) \\
&\geq (1 - \nu_K) \mathbb{P}(T_0^{(1,+)} < \infty)
\end{aligned}$$

where  $\nu_K$  is a sequence converging to zero as  $K \rightarrow \infty$ . for large enough  $K$ . Thus,

$$\begin{aligned}
\mathbb{P}(T_{\varepsilon K}^{(1)} < \tilde{T}_\varepsilon^{(0)}) &= 1 - \mathbb{P}(\tilde{T}_\varepsilon^{(0)} < T_{\varepsilon K}^{(1)}) \leq 1 - \mathbb{P}(T_0^{(1)} < \tilde{T}_\varepsilon^{(0)} < T_{\varepsilon K}^{(1)}) \\
&\leq 1 - (1 - \nu_K) \mathbb{P}(T_0^{(1,+)} < \infty) \rightarrow 1 - \frac{d_1^+}{b_1^+}.
\end{aligned} \tag{4.14}$$

Combining (4.13) and (4.14), there exists  $\beta > 0$  such that for large enough  $K$ ,

$$\begin{aligned}
& \mathbb{P}\left((1 - c\varepsilon)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1^+} < T_{\varepsilon K}^{(1)} < (1 + c\varepsilon)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1^+} \mid T_{\varepsilon K}^{(1)} < \tilde{T}_\varepsilon^{(0)}\right) \\
&= \frac{\mathbb{P}\left((1 - c\varepsilon)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1^+} < T_{\varepsilon K}^{(1)} < (1 + c\varepsilon)\frac{\log K}{\frac{1}{2}(b_1^- + b_1^+) - d_1^+}, T_{\varepsilon K}^{(1)} < \tilde{T}_\varepsilon^{(0)}\right)}{\mathbb{P}(T_{\varepsilon K}^{(1)} < \tilde{T}_\varepsilon^{(0)})} \geq 1 - \beta\varepsilon.
\end{aligned}$$

□

## 4.4 Second Phase

In the second phase, once the type 1 population has reached size  $\varepsilon K$  for some  $\varepsilon > 0$ , it is known (see [5]) in the static case that the paths of the coupled logistic birth-death process normalised by  $K$  stay within  $\varepsilon$  of the approximating deterministic system for time of order  $O(1)$  with probability approaching 1 as  $K \rightarrow \infty$ . Since the deterministic system (a competitive Lotka-Volterra system in this case) reaches an  $\varepsilon$ -neighbourhood of the stable equilibrium of the system in  $O(1)$  time, in the periodic case with  $S_K \gg O(1)$ , the normalised population sizes  $(\frac{N_0(t)}{K}, \frac{N_1(t)}{K})_{t \geq 0}$  will simply follow the paths of the deterministic system with parameters corresponding to the period in which  $T_{\varepsilon K}^{(1)}$  falls.

If  $S_K \rightarrow 0$ , an averaging will occur in which the normalised population sizes will approximate the averaged deterministic system. This is made precise in the next proposition, which follows a similar proof structure to Theorem 2.1 in Chapter 11 of [17].

**Proposition 4.4.1.** *Fix  $T > 0$ . Consider a pair of coupled, time-dependent logistic birth-death processes  $\mathbf{N}(t, \mathbf{y}) = (N_0(t), N_1(t))_{t \geq 0}$  with initial condition  $\mathbf{N}(0, \mathbf{y}) = (y_0 K, y_1 K)$  for some  $\mathbf{y} = (y_0, y_1) \in (0, \infty)^2$ , such that the type 0 birth and death rates are fixed and the type 1 birth and death rates are determined by periodic and piecewise constant functions of period  $2S_K$  such that  $S_K \rightarrow 0$  as  $K \rightarrow \infty$ . Let  $b_K(s)$  and  $d_K(s)$  denote the birth and death rates of the type 1 population taking values  $(b_+, d_+)$  and  $(b_-, d_-)$  in odd and even periods respectively, with  $b_- - d_- < b_+ - d_+$ , and let  $b_0$  and  $d_0$  denote the birth and death rates of the type 0 population. Suppose also that the competition parameters  $C_{0,0}$ ,  $C_{0,1}$  and  $C_{1,1}$  are fixed. Define the averaged rates of the type 1 population,  $\bar{b} = \frac{1}{2}(b_+ + b_-)$  and  $\bar{d} = \frac{1}{2}(d_+ + d_-)$ , and for  $t \geq 0$  let  $\mathbf{X}(t, \mathbf{y}) = (X_0(t, y_0), X_1(t, y_1))$  denote the solution to the deterministic system of ODEs*

$$\begin{aligned} \frac{dX_0}{dt} &= (b_0 - d_0)X_0 - C_{0,1}X_0X_1 - C_{0,0}X_0^2 \\ \frac{dX_1}{dt} &= (\bar{b} - \bar{d})X_1 - C_{0,1}X_0X_1 - C_{1,1}X_1^2 \end{aligned} \tag{4.15}$$

subject to the initial condition  $(X_0(0), X_1(0)) = (y_0, y_1)$ . Then for all compact sets  $D \subseteq [0, \infty)$  and  $\varepsilon > 0$ ,

$$\sup_{(y_0, y_1) \in D} \mathbb{P} \left( \sup_{t \in [0, T]} \left\| \frac{1}{K} \mathbf{N}(t, \mathbf{y}) - \mathbf{X}(t, \mathbf{y}) \right\|_1 > \varepsilon \right) \rightarrow 0$$

as  $K \rightarrow \infty$ .

*Proof.* Let  $P_b(t)$  and  $P_d(t)$  be independent, rate 1 Poisson processes. Then for all  $t \geq 0$  we can write

$$N_0(t) = N_0(0) + P_b \left( b_0 \int_0^t N_0(s) ds \right) - P_d \left( \int_0^t \left( d_0 + \frac{C_{0,0}N_0(s)}{K} + \frac{C_{0,1}N_1(s)}{K} \right) N_0(s) ds \right)$$

and

$$\begin{aligned} N_1(t) &= N_1(0) + P_b \left( \int_0^t b \left( \frac{s}{S_K} \right) N_1(s) ds \right) \\ &\quad - P_d \left( \int_0^t \left( d \left( \frac{s}{S_K} \right) + \frac{C_{0,1}N_0(s)}{K} + \frac{C_{1,1}N_1(s)}{K} \right) N_1(s) ds \right) \end{aligned}$$

where  $b(t)$  and  $d(t)$  are functions of period 2 such that  $b_K((2i+r)S_K) = b(r)$  and  $d_K((2i+r)S_K) = d(r)$  for all  $i \in \mathbb{N} \cup \{0\}$  and  $r \in [0, 2)$ . Defining  $X_0^K(t) = \frac{N_0(t)}{K}$  and  $X_1^K(t) = \frac{N_1(t)}{K}$  we have

$$\begin{aligned} X_0^K(t) &= X_0(0) + \frac{1}{K} P_b \left( K b_0 \int_0^t X_0^K(s) ds \right) \\ &\quad - \frac{1}{K} P_d \left( K \int_0^t \left( d_0 + C_{0,0}X_0^K(s) + C_{0,1}X_1^K(s) \right) X_0^K(s) ds \right) \end{aligned}$$

and

$$\begin{aligned} X_1^K(t) &= X_1(0) + \frac{1}{K} P_b \left( K \int_0^t b \left( \frac{s}{S_K} \right) X_1^K(s) ds \right) \\ &\quad - \frac{1}{K} P_d \left( K \int_0^t \left( d \left( \frac{s}{S_K} \right) + C_{0,1}X_0^K(s) + C_{1,1}X_1^K(s) \right) X_1^K(s) ds \right). \end{aligned}$$

Recall that  $M_b(t) = P_b(t) - t$  and  $M_d(t) = P_d(t) - t$  are martingales for which we can compute the predictable quadratic variation. By calculation, we have

$$\langle M_b \rangle_t = \langle M_d \rangle_t = t.$$

Now define the quantities

$$\delta_0^K(t, y_0) = \left| X_0^K(t) - X_0^K(0) - \int_0^t \left( (b_0 - d_0)X_0^K(s) - C_{0,1}X_1^K(s)X_0^K(s) - C_{0,0}X_0^K(s)^2 \right) ds \right|$$

and

$$\delta_1^K(t, y_1) = \left| X_1^K(t) - X_1^K(0) - \int_0^t \left( (\bar{b} - \bar{d})X_1^K(s) - C_{0,1}X_1^K(s)X_0^K(s) - C_{1,1}X_1^K(s)^2 \right) ds \right|.$$

Additionally define

$$\begin{aligned}\tilde{\delta}_1^K(t, y_1) = & \left| X_1^K(t) - X_1^K(0) - \int_0^t \left( b\left(\frac{s}{S_K}\right) - d\left(\frac{s}{S_K}\right) \right) X_1^K(s) ds \right. \\ & \left. + \int_0^t \left( C_{0,1} X_1^K(s) X_0^K(s) + C_{1,1} X_1^K(s)^2 \right) ds \right|\end{aligned}$$

so that

$$\delta_1^K(t, y_1) = \tilde{\delta}_1^K(t, y_1) + \left| \int_0^t \left( \left( b\left(\frac{s}{S_K}\right) - \bar{b} \right) - \left( d\left(\frac{s}{S_K}\right) - \bar{d} \right) \right) X_1^K(s) ds \right|.$$

Set  $\varepsilon_0^K(T, y_0) = \sup_{t \in [0, T]} \delta_0^K(t, y_0)$  and  $\tilde{\varepsilon}_1^K(T, y_1) = \sup_{t \in [0, T]} \tilde{\delta}_1^K(t, y_1)$ . Observe that

$$\begin{aligned}X_0^K(t) = & X_0^K(0) + \frac{1}{K} M_b \left( K b_0 \int_0^t X_0^K(s) ds \right) + b_0 \int_0^t X_0^K(s) ds \\ & - \frac{1}{K} M_d \left( K \int_0^t \left( d_0 + C_{0,0} X_0^K(s) + C_{0,1} X_1^K(s) \right) X_0^K(s) ds \right) \\ & - \int_0^t \left( d_0 + C_{0,0} X_0^K(s) + C_{0,1} X_1^K(s) \right) X_0^K(s) ds.\end{aligned}$$

and

$$\begin{aligned}X_1^K(t) = & X_1^K(0) + \frac{1}{K} M_b \left( K \int_0^t b\left(\frac{s}{S_K}\right) X_1^K(s) ds \right) + \int_0^t b\left(\frac{s}{S_K}\right) X_1^K(s) ds \\ & - \frac{1}{K} M_d \left( K \int_0^t \left( d\left(\frac{s}{S_K}\right) + C_{0,1} X_0^K(s) + C_{1,1} X_1^K(s) \right) X_1^K(s) ds \right) \\ & - \int_0^t \left( d\left(\frac{s}{S_K}\right) + C_{0,1} X_0^K(s) + C_{1,1} X_1^K(s) \right) X_1^K(s) ds.\end{aligned}$$

From the above representations we have

$$\begin{aligned}\tilde{\varepsilon}_1^K(T, y_1) = & \sup_{t \in [0, T]} \left| \frac{1}{K} M_b \left( K \int_0^t b\left(\frac{s}{S_K}\right) X_1^K(s) ds \right) \right. \\ & \left. - \frac{1}{K} M_d \left( K \int_0^t \left( d\left(\frac{s}{S_K}\right) + C_{0,1} X_0^K(s) + C_{1,1} X_1^K(s) \right) X_1^K(s) ds \right) \right|\end{aligned}$$

and

$$\begin{aligned}\varepsilon_0^K(T, y_0) = & \sup_{t \in [0, T]} \left| \frac{1}{K} M_b \left( K b_0 \int_0^t X_0^K(s) ds \right) \right. \\ & \left. - \frac{1}{K} M_d \left( K \int_0^t \left( d_0 + C_{0,0} X_0^K(s) + C_{0,1} X_1^K(s) \right) X_0^K(s) ds \right) \right|.\end{aligned}$$



For each  $K \geq 1$ , the intrinsic birth rate of the process  $N_1(t)$  in any period never exceeds  $b_+$  and the process  $N_0(t)$  has constant birth rate  $b_0$ , so defining  $b_{\max} = b_+ \vee b_0$  we can construct the process

$$\hat{N}(t) = N_0(0) \vee N_1(0) + P_b \left( \int_0^t b_{\max} \hat{N}(s) ds \right). \quad (4.16)$$

This is a standard (linear) pure birth process with intrinsic birth rate  $b_{\max}$  whose only dependence on  $K$  comes through the initial condition  $N_0(0) \vee N_1(0) = (y_0 \vee y_1)K$ . Also, since we used the same rate 1 Poisson process  $P_b$  in the definition of the processes  $N_0(t)$  and  $N_1(t)$ , we have the coupling

$$N_0(t) \vee N_1(t) \leq \hat{N}(t) \quad \text{for all } t \geq 0,$$

which holds for all  $K \geq 1$ , almost surely. Defining  $\hat{X}^K(t) = \frac{\hat{N}(t)}{K}$ , it follows that for each  $K \geq 1$ ,

$$\sup_{t \in [0, T]} \{X_0^K(t) \vee X_1^K(t)\} \leq \sup_{t \in [0, T]} \hat{X}^K(t)$$

almost surely. Note also that since  $\hat{N}(t)$  is a pure birth process,  $\sup_{t \in [0, T]} \hat{N}(t) = \hat{N}(T)$

and from the representation (4.16) it follows that  $\mathbb{E}(\hat{N}(t)) = \hat{N}(0)e^{b_{\max}t}$  for all  $t \geq 0$ . Hence for each  $K \geq 1$  we have,

$$\mathbb{E} \left( \sup_{t \in [0, T]} \hat{N}(t) \right) = \mathbb{E}(\hat{N}(T)) = (y_0 \vee y_1)K e^{b_{\max}T}$$

which implies

$$\sup_{K \geq 1} \mathbb{E} \left( \sup_{t \in [0, T]} \hat{X}^K(t) \right) < \infty. \quad (4.17)$$

For some  $\alpha \in (0, \frac{1}{2})$  and each  $K \geq 1$ , define  $R'_K = K^{\frac{1}{2}-\alpha}$ . Defining the event  $B_K = \{ \sup_{t \in [0, T]} X_0^K(t) \vee X_1^K(t) \leq R'_K \}$ , we have by Markov's inequality and (4.17),

$$\mathbb{P}(B_K^c) \leq \left( \frac{1}{R'_K} \right) \mathbb{E} \left( \sup_{t \in [0, T]} \hat{X}^K(t) \right) \leq \left( \frac{1}{R'_K} \right) \sup_{K \geq 1} \mathbb{E} \left( \sup_{t \in [0, T]} \hat{X}^K(t) \right) \leq V' K^{\alpha-\frac{1}{2}}$$

for some constant  $V' > 0$ . Thus  $\mathbb{P}(B_K^c) \rightarrow 0$  as  $K \rightarrow \infty$ . Note that the functions  $f_b^0(x) = b_0x$ ,  $f_b^1(x) = b_1^+x$ ,  $f_d^0(x_0, x_1) = (d_0 + C_{0,0}x_0 + C_{0,1}x_1)x_0$  and  $f_d^1(x_0, x_1) = (d_1^- + C_{0,1}x_0 + C_{1,1}x_1)x_1$  are continuous and hence bounded on any compact subsets

of  $[0, \infty)$  and  $[0, \infty)^2$ , respectively. Thus we can define for each  $K \geq 1$ ,

$$R_K = \max \left\{ \sup_{x \in [0, R'_K]} f_b^0(x), \sup_{x \in [0, R'_K]} f_b^1(x), \sup_{(x_0, x_1) \in [0, R'_K]^2} f_d^0(x_0, x_1), \sup_{(x_0, x_1) \in [0, R'_K]^2} f_d^1(x_0, x_1) \right\}.$$

Since the functions in the above maximum have at most quadratic growth, there exists a constant  $V > 0$  such that for large enough  $K$ ,

$$R_K \leq V(R'_K)^2 = VK^{1-2\alpha}.$$

The aim is to show that for all  $\varepsilon > 0$ ,  $\mathbb{P}(\tilde{\varepsilon}_1^K(T, y_1) > \varepsilon) \rightarrow 0$  and  $\mathbb{P}(\varepsilon_0^K(T, y_0) > \varepsilon) \rightarrow 0$  as  $K \rightarrow \infty$ . First note that

$$\mathbb{P}(\tilde{\varepsilon}_1^K(T, y_1) > \varepsilon) \leq \mathbb{P}(\tilde{\varepsilon}_1^K(T, y_1) > \varepsilon | B_K) + \mathbb{P}(B_K^c). \quad (4.18)$$

On the event  $B_K$  we have

$$\tilde{\varepsilon}_1^K(T, y_1) \leq \frac{1}{K} \sup_{u \in [0, R_K T]} |M_b(Ku)| + \frac{1}{K} \sup_{u \in [0, R_K T]} |M_d(Ku)| \quad (4.19)$$

Let  $\langle M_b \rangle_t = t$  denote the predictable quadratic variation of  $M_b$  and recall that  $W(t) = M_b(t)^2 - \langle M_b \rangle_t$  is a martingale starting at zero. Hence,

$$\mathbb{E}(M_b(t)^2) = \mathbb{E}(W(t)) + \mathbb{E}(\langle M_b \rangle_t) = t.$$

By Doob's martingale inequality we therefore have

$$\mathbb{P}\left(\frac{1}{K} \sup_{u \in [0, R_K T]} |M_b(Ku)| > \frac{\varepsilon}{2}\right) \leq \frac{4}{\varepsilon^2 K^2} \mathbb{E}(M_b(K R_K T)^2) = \frac{4 R_K T}{\varepsilon^2 K} \leq \left(\frac{4VT}{\varepsilon^2}\right) K^{-2\alpha} \quad (4.20)$$

and in the same way

$$\mathbb{P}\left(\frac{1}{K} \sup_{u \in [0, R_K T]} |M_d(Ku)| > \frac{\varepsilon}{2}\right) \leq \frac{4 R_K T}{\varepsilon^2 K} \leq \left(\frac{4VT}{\varepsilon^2}\right) K^{-2\alpha}. \quad (4.21)$$

From (4.19) we then have  $\mathbb{P}(\tilde{\varepsilon}_1^K(T, y_1) > \varepsilon | B_K) \rightarrow 0$  as  $K \rightarrow \infty$ . From (4.18) and the fact that  $\mathbb{P}(B_K^c) \rightarrow 0$  which we showed earlier,

$$\mathbb{P}(\tilde{\varepsilon}_1^K(T, y_1) > \varepsilon) \rightarrow 0 \quad (4.22)$$

as  $K \rightarrow \infty$ . Applying the same strategy shows that

$$\mathbb{P}(\varepsilon_0^K(T, y_0) > \varepsilon) \rightarrow 0 \quad (4.23)$$

as  $K \rightarrow \infty$ .

As the function  $g_0(x, y) = (b_0 - d_0)x - C_{0,1}xy - C_{0,0}x^2$  is locally Lipschitz continuous in  $x$ , for any compact set  $D \subseteq [0, \infty)$  we have a minimal Lipschitz constant  $L_{0,y}$  for each  $y$ , given by

$$L_{0,y} = \sup_{x_1, x_2 \in D} \frac{|g_0(x_2, y) - g_0(x_1, y)|}{|x_2 - x_1|}.$$

Since  $g_0(x, y)$  is continuous in both variables,  $L_{0,y}$  is continuous as a function of  $y$  on any compact set. Thus on any compact set  $E \subseteq [0, \infty)$  one can choose a maximal Lipschitz constant  $L_0 = \sup_{y \in E} L_{0,y}$  that holds for  $g_0(\cdot, y)$  for all  $y \in E$ . Similarly, such a constant  $L_1 > 0$  exists for the function  $g_1(x, y) = (\bar{b} - \bar{d})x - C_{0,1}xy - C_{1,1}x^2$ . Since the solutions of the ODE system (4.15) are bounded on any finite time interval, we have

$$\begin{aligned} |X_0^K(t) - X_0(t)| &\leq \delta_0^K(t, y_0) + L_0 \int_0^t |X_0^K(s) - X_0(s)| ds \\ |X_1^K(t) - X_1(t)| &\leq \delta_1^K(t, y_1) + L_1 \int_0^t |X_1^K(s) - X_1(s)| ds \end{aligned}$$

for all  $t \in [0, T]$  almost surely. Recall that we defined  $\tilde{\delta}_1^K(t, y_1)$  such that

$$\delta_1^K(t, y_1) = \tilde{\delta}_1^K(t, y_1) + \left| \int_0^t \left( \left( b\left(\frac{s}{S_K}\right) - \bar{b} \right) - \left( d\left(\frac{s}{S_K}\right) - \bar{d} \right) \right) X_1^K(s) ds \right|.$$

We aim to bound the second term above for large  $K$ . Define

$$\eta_1^K(t, y_1) = \left| \int_0^t \left( \left( b\left(\frac{s}{S_K}\right) - \bar{d} \right) - \left( d\left(\frac{s}{S_K}\right) - \bar{d} \right) \right) X_1(s) ds \right|.$$

Note that

$$\begin{aligned}
& \left| \int_0^t \left( \left( b\left(\frac{s}{S_K}\right) - \bar{b} \right) - \left( d\left(\frac{s}{S_K}\right) - \bar{d} \right) \right) X_1^K(s) ds \right| \\
& \leq \left| \int_0^t \left( \left( b\left(\frac{s}{S_K}\right) - \bar{b} \right) - \left( d\left(\frac{s}{S_K}\right) - \bar{d} \right) \right) (X_K(s) - X(s)) ds \right| \\
& + \left| \int_0^t \left( \left( b\left(\frac{s}{S_K}\right) - \bar{b} \right) - \left( d\left(\frac{s}{S_K}\right) - \bar{d} \right) \right) X_1^K(s) ds \right| \\
& \leq ((b_+ - d_+) - (\bar{b} - \bar{d})) \int_0^t |X_1^K(s) - X_1(s)| ds \\
& + \eta_1^K(t, y_1).
\end{aligned}$$

Defining the sequences  $m_K \in \mathbb{N}$  and  $r_K \in [0, 2)$  such that  $t = 2m_K S_K + r_K S_K$  for each  $K \in \mathbb{N}$ , we have

$$\begin{aligned}
\eta_1^K(t, y_1) &= \left| \int_0^t \left( \left( b\left(\frac{s}{S_K}\right) - \bar{b} \right) - \left( \mu\left(\frac{s}{S_K}\right) - \bar{d} \right) \right) X_1(s) ds \right| \\
&\leq \sum_{i=1}^{m_K} \left| \int_{2(i-1)S_K}^{2iS_K} \left( \left( b\left(\frac{s}{S_K}\right) - \bar{b} \right) - \left( d\left(\frac{s}{S_K}\right) - \bar{d} \right) \right) X_1(s) ds \right| \\
&+ \left| \int_{2m_K S_K}^{2m_K S_K + r_K S_K} \left( \left( b\left(\frac{s}{S_K}\right) - \bar{b} \right) - \left( d\left(\frac{s}{S_K}\right) - \bar{d} \right) \right) X_1(s) ds \right|
\end{aligned}$$

Note that the solutions to the ODE system (4.15) are continuously differentiable, and hence locally Lipschitz. Using this fact and applying Lemma 4.6.3 (see Auxiliary Results), the integrals in the sum are given by

$$\begin{aligned}
& \int_{2(i-1)S_K}^{2iS_K} \left( \left( b\left(\frac{s}{S_K}\right) - \bar{b} \right) - \left( d\left(\frac{s}{S_K}\right) - \bar{d} \right) \right) X_1(s) ds \\
&= ((b_+ - d_+) - (\bar{b} - \bar{d})) \int_{2(i-1)S_K}^{2iS_K} X_1(s) ds + ((b_- - d_-) - (\bar{b} - \bar{d})) \int_{2(i-1)S_K}^{2iS_K} X_1(s) ds \\
&= ((b_+ - d_+) - (\bar{b} - \bar{d})) \left( \int_{2(i-1)S_K}^{2iS_K} X_1(s) ds - \int_{2(i-1)S_K}^{2iS_K} X_1(s) ds \right) \\
&\leq ((b_+ - d_+) - (\bar{b} - \bar{d})) Q_1 S_K^2
\end{aligned}$$

for a constant  $Q_1 > 0$ , where we used the fact that the function  $X_1$  is Lipschitz continuous on  $[0, T]$ . Additionally by boundedness of  $X_1$ ,

$$\left| \int_{2m_K S_K}^{2m_K S_K + r_K S_K} \left( \left( b\left(\frac{s}{S_K}\right) - \bar{b} \right) - \left( d\left(\frac{s}{S_K}\right) - \bar{d} \right) \right) X_1(s) ds \right| \leq C((b_+ - d_+) - (\bar{b} - \bar{d})) S_K.$$

for some  $C > 0$ . Therefore

$$\begin{aligned}\eta_1^K(t, y_1) &\leq ((b_+ - d_+) - (\bar{b} - \bar{d}))Q_1 m_K S_K^2 + C((b_+ - d_+) - (\bar{b} - \bar{d}))S_K \\ &\leq C'((b_+ - d_+) - (\bar{b} - \bar{d}))tS_K\end{aligned}$$

for some  $C' > 0$ . Thus we have almost surely that

$$\begin{aligned}|X_1^K(t) - X_1(t)| &\leq \left( \tilde{\delta}_1^K(t, y_1) + \eta_1^K(t, y_1) \right) \\ &\quad + (L_1 + ((b_+ - d_+) - (\bar{b} - \bar{d}))) \int_0^t |X_1^K(s) - X_1(s)| ds\end{aligned}$$

Recall from before we also have that

$$|X_0^K(t) - X_0(t)| \leq \delta_0^K(t, y_0) + L_0 \int_0^t |X_0^K(s) - X_0(s)| ds$$

almost surely. By Gronwall's inequality, the following hold almost surely:

$$\begin{aligned}|X_0^K(t) - X_0(t)| &\leq \delta_0^K(t, y_0)e^{L_0 t} \implies \sup_{t \in [0, T]} |X_0^K(t) - X_0(t)| \leq \varepsilon_0^K(T, y_0)e^{L_0 T} \\ |X_1^K(t) - X_1(t)| &\leq (\tilde{\delta}_1^K(t, y_1) + \eta_1^K(t, y_1))e^{(L_1 + ((b_+ - d_+) - (\bar{b} - \bar{d})))t} \\ &\implies \sup_{t \in [0, T]} |X_1^K(t) - X_1(t)| \leq (\tilde{\varepsilon}_1^K(T, y_1) + \eta_1^K(T, y_1))e^{(L_1 + ((b_+ - d_+) - (\bar{b} - \bar{d})))T}.\end{aligned}$$

Define  $\mathbf{y} = (y_0, y_1)$  and  $\mathbf{X}_K(t, \mathbf{y}) = (X_0^K(t, y_0), X_1^K(t, y_1))$ ,  $\mathbf{X}(t, \mathbf{y}) = (X_0(t, y_0), X_1(t, y_1))$ .

We wish to determine, for all  $\varepsilon > 0$  and compact sets  $D \subseteq [0, \infty)^2$ , the limit

$$\lim_{K \rightarrow \infty} \sup_{(y_0, y_1) \in D} \mathbb{P} \left( \sup_{t \in [0, T]} \|\mathbf{X}_K(t, \mathbf{y}) - \mathbf{X}(t, \mathbf{y})\|_1 > \varepsilon \right).$$

Noting (4.22) and (4.23), and using the assumption  $S_K \rightarrow 0$  we have

$$\begin{aligned}
& \sup_{(y_0, y_1) \in D} \mathbb{P} \left( \sup_{t \in [0, T]} \|\mathbf{X}_K(t, \mathbf{y}) - \mathbf{X}(t, \mathbf{y})\|_1 > \varepsilon \right) \\
& \leq \sup_{(y_0, y_1) \in D} \mathbb{P} \left( \sup_{t \in [0, T]} |X_0^K(t) - X_0(t)| > \frac{\varepsilon}{2} \right) + \sup_{(y_0, y_1) \in D} \mathbb{P} \left( \sup_{t \in [0, T]} |X_1^K(t) - X_1(t)| > \frac{\varepsilon}{2} \right) \\
& \leq \sup_{(y_0, y_1) \in D} \mathbb{P} \left( \tilde{\varepsilon}_1^K(T, y_1) > \frac{e^{-(L_1 + ((b_+ - d_+) - (\bar{b} - \bar{d})))T} \varepsilon}{4} \right) \\
& + \sup_{(y_0, y_1) \in D} \mathbb{P} \left( \eta_1^K(T, y_1) > \frac{e^{-(L_1 + ((b_+ - d_+) - (\bar{b} - \bar{d})))T} \varepsilon}{4} \right) \\
& + \sup_{(y_0, y_1) \in D} \mathbb{P} \left( \varepsilon_0^K(T, y_0) > \frac{e^{-L_0 T} \varepsilon}{2} \right) \\
& \rightarrow 0
\end{aligned}$$

as  $K \rightarrow \infty$ . □

## 4.5 Possible Extension to Random Environment

A natural extension of Theorem 4.3.2 is to suppose that the type 1 process has birth and death rates  $b_K$  and  $d_K$  which themselves evolve as a random processes in time according to an independent Poisson process, as opposed to a deterministic periodic variation. This is in a small way analogous to models of population genetics for fixed-size populations in which the fitness of a given allele can take values from a set  $A$  from which the fitness is uniformly sampled at the arrival times of a Poisson process. A Moran model with fluctuating selection of this type is studied in [6], where the key result relies on a theorem of Kurtz [26] relating to stochastic averaging which we state below.

To state the theorem requires the following definitions: let  $(S, d)$  be a metric space and let  $\mathcal{M}([0, \infty) \times S)$  denote the space of finite Borel measures on  $[0, \infty) \times S$ . For each  $t \geq 0$ , let  $d_t$  denote the product metric on  $[0, t] \times S$  and let  $\mathcal{M}([0, t] \times S)$  denote the set of finite Borel measures on  $[0, t] \times S$  equipped with the weak topology<sup>2</sup>. The weak topology is metrized by the *Prokhorov metric*  $\rho_t$ , which for two measures  $\mu, \nu \in \mathcal{M}([0, t] \times S)$  is

---

<sup>2</sup>A sequence of such finite measures  $(\mu_n)_n$  is said to converge to a measure  $\mu$  with respect to the weak topology if for all bounded, real-valued continuous functions  $f : X \rightarrow \mathbb{R}$  we have  $\int_X f(x) \mu_n(dx) \rightarrow \int_X f(x) \mu(dx)$ .

given by

$$\rho_t(\mu, \nu) = \inf \left\{ \varepsilon > 0 : \mu(A) \leq \nu(A^\varepsilon) + \varepsilon, \nu(A) \leq \mu(A^\varepsilon) + \varepsilon \text{ for all } A \in \mathcal{B}(\mathcal{M}([0, t] \times S)) \right\}$$

where  $\mathcal{B}(\mathcal{M}([0, t] \times S))$  is the Borel  $\sigma$ -algebra on  $\mathcal{M}([0, t] \times S)$  and  $A^\varepsilon = \{y \in \mathcal{M}([0, t] \times S) : d_t(x, y) < \varepsilon \text{ for some } x \in A\}$ . For each  $\mu \in \mathcal{M}([0, \infty) \times S)$ , let  $\mu_t$  denote the restriction of  $\mu$  to the space  $\mathcal{M}([0, t] \times S)$  and specify  $l_m(S) \subseteq \mathcal{M}([0, \infty) \times S)$  to contain all measures  $\mu \in \mathcal{M}([0, \infty) \times S)$  such that  $\mu([0, t] \times S) = t$  for each  $t \geq 0$ . Define the metric  $\hat{\rho}$  for measures  $\mu, \nu \in l_m(S)$  by

$$\hat{\rho}(\mu, \nu) = \int_0^\infty e^{-t} (1 \wedge \rho_t(\mu_t, \nu_t)) dt$$

and note that for a sequence  $(\mu^n)_n \in l_m(S)$  and  $\mu \in l_m(S)$ ,  $\hat{\rho}(\mu^n, \mu) \rightarrow 0$  if and only if  $\rho_t(\mu_t^n, \mu_t) \rightarrow 0$  for almost every  $t \geq 0$ . From now on the notation  $l_m(S)$  refers to the metric space  $(l_m(S), \hat{\rho})$ .

**Theorem 4.5.1** (Theorem 2.1 in [26]). *Let  $(E_1, d_1)$  and  $(E_2, d_2)$  be complete, separable metric spaces and set  $E = E_1 \times E_2$ . Let  $(X_n, Y_n)_n$  be a sequence of stochastic processes with sample paths in  $D_E[0, \infty)$  such that  $(X_n(t), Y_n(t))_{t \geq 0}$  is adapted to a filtration  $\mathcal{F}_n = (\mathcal{F}_t^n)_{t \geq 0}$  for each  $n \geq 1$ . Suppose  $X_n$  satisfies the compact containment condition: for each  $\varepsilon > 0$  and  $T > 0$  there exists a compact set  $K \subseteq E_1$  such that*

$$\inf_{n \geq 1} \mathbb{P} \left( \bigcap_{t \in [0, T]} \{X_n(t) \in K\} \right) \geq 1 - \varepsilon$$

and assume that the set  $\cup_{n=1}^\infty \cup_{t \in [0, \infty)} \{Y_n(t)\}$  has compact closure in  $E_2$ . Let  $\mathcal{D}(A) \subseteq C_b(E_1)$  be the domain of an operator  $A : \mathcal{D}(A) \rightarrow C(E_1 \times E_2)$  such that  $\mathcal{D}(A)$  is dense in  $C_b(E_1)$  with respect to the compact-open topology, and such that for each  $f \in \mathcal{D}(A)$ , there exists a stochastic process  $(\varepsilon_n^f(t))_{t \geq 0}$  for which, defining  $g_f = Af$ ,

$$M_n^f(t) = f(X_n(t)) - \int_0^t g_f(X_n(s), Y_n(s)) ds + \varepsilon_n^f(t)$$

is an  $\mathcal{F}_t^n$ -martingale. Additionally suppose that the following two conditions hold: for each  $f \in \mathcal{D}(A)$  and each  $T > 0$ ,

(i) *There exists  $p > 1$  such that*

$$\sup_{n \geq 1} \mathbb{E} \left( \int_0^T |g_f(X_n(t), Y_n(t))|^p dt \right) < \infty \quad (4.24)$$

(ii) The process  $(\varepsilon_n^f(t))_{t \geq 0}$  satisfies

$$\lim_{n \rightarrow \infty} \mathbb{E} \left( \sup_{t \in [0, T]} |\varepsilon_n^f(t)| \right) = 0 \quad (4.25)$$

Let  $(\Gamma_n)_n$  be a sequence of random variables taking values in  $l_m(E_2)$  given by

$$M^f(t) = \Gamma_n([0, t] \times B) = \int_0^t \mathbb{1}_B(Y_n(s)) ds. \quad (4.26)$$

Then the sequence  $(X_n, \Gamma_n)_n$  has compact closure in  $D_{E_1}[0, \infty) \times l_m(E_2)$  and for any limit point  $(X, \Gamma)$  there exists a filtration  $(\mathcal{G}_t)_t$  such that for each  $f \in \mathcal{D}(A)$ ,

$$f(X(t)) - \int_0^t \int_{E_2} g_f(X(s), y) \Gamma(ds, dy) \quad (4.27)$$

is a  $\mathcal{G}_t$ -martingale.

To apply the above theorem, one could choose  $b_+, d_+, b_-, d_- \in (0, \infty)$  such that  $b_+ + d_+ = b_- + d_- = a$  for some positive constant  $a$ , and  $y_- = b_- - d_- < b_+ - d_+ = y_+$ . Define the set  $E_Y = \{y_-, y_+\}$ . Fixing the sum of rates to equal a constant ensures that for any value of  $y \in E_Y$ , the birth and death rates can be recovered via  $b(y) = \frac{1}{2}(a + y)$  and  $d(y) = \frac{1}{2}(a - y)$ . Now let  $(Y_K(t))_{t \geq 0}$  denote a marked Poisson point process of rate  $\lambda_K > 0$  on  $E_Y$ , which means at rate  $\lambda_K$ , the value of  $Y_K$  is uniformly resampled from the set  $E_Y$ . A logistic birth-death process  $(N_K(t))_{t \geq 0}$  in random environment  $Y_K$  can then be defined as follows: set  $N_K(0) = \varepsilon K$  for some  $\varepsilon > 0$ , and let  $(Y_K(t))_{t \geq 0}$  be the marked Poisson point process defined above. The process  $(N_K(t))_{t \geq 0}$  then evolves in  $\mathbb{N}$  by gaining and losing individuals according to two dependent Cox processes - a birth intensity process  $B_K(t)$  and death intensity process  $D_K(t)$ , such that the birth intensity at time  $t$  is  $b(Y_K(t))$  and the death intensity is  $d(Y_K(t)) + \frac{CN_K(t)}{K}$ , where  $C > 0$  is a competition parameter. Note that the stochastic process  $(N_K(t))_{t \geq 0}$  is not a Markov process, however the environment process  $(Y_K(t))_{t \geq 0}$  and the joint process  $(N_K(t), Y_K(t))_{t \geq 0}$  are Markov processes. Thus, defining  $X_K(t) = \frac{N_K(t)}{K}$ , the process  $(X_K(t), Y_K(t))_{t \geq 0}$  is also a Markov process, and in particular can be characterised as the solution to a martingale problem for its infinitesimal generator. However as the non-Markov process  $(X_K(t))_{t \geq 0}$  is fundamentally of interest, the aim is to construct an ‘approximate martingale problem’ for  $X_K$  and use Theorem 4.4.2 above to show that in the limit  $K \rightarrow \infty$ , the dependence on  $Y_K$  ‘averages out’ to yield a true martingale problem, the solution to which is a time-homogeneous Markov process  $(X(t))_{t \geq 0}$  on the same state space as  $(X_K(t))_{t \geq 0}$ . In particular, the process  $Y(tK) = Y_K(t)$  approaches



its stationary distribution as  $K \rightarrow \infty$ , which is the uniform distribution on the set  $E_Y$ . Thus the limiting measure  $\Gamma$  in equation (4.27) above is equal to the stationary distribution of the continuous-time Markov chain  $Y$ , since  $\Gamma$  is the weak limit of the occupation measures  $(\Gamma_n)_n$  in equation (4.26).

## 4.6 Auxiliary Results

**Lemma 4.6.1.** *Let  $(N(t))_{t \geq 0}$  be a time-varying birth-death process with right-continuous, bounded and piecewise constant birth and death intensities,  $b(t)$  and  $d(t)$  respectively, and let  $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$  be its natural filtration. Define the function  $\rho(t) = \int_0^t (d(s) - b(s)) ds$  and suppose that  $\int_0^\infty e^{\rho(t)} dt < \infty$ . Then the process  $M(t) = N(t)e^{\rho(t)}$  is a square integrable  $\mathcal{F}_t$ -martingale.*

*Proof.* First note that the process  $(N(t))_{t \geq 0}$  has a family of right-generators  $(A_t^+)_{t \geq 0}$  which are operators mapping the space of bounded, measurable functions into itself, given by

$$A_t^+ f(n) = \lim_{s \rightarrow 0} \frac{\mathbb{E}(f(N(t+s)) | N(t) = n) - f(n)}{s}$$

for all  $f \in \mathcal{D}(A^+)$ , where  $\mathcal{D}(A^+)$  is the set of measurable functions  $f : \mathbb{N} \cup \{0\} \rightarrow \mathbb{R}$  for which the above limit exists for all  $n \in \mathbb{N} \cup \{0\}$  and  $t \geq 0$ . It is then a known result (see Theorem 2.1 in [27]) that for any  $f \in \mathcal{D}(A^+)$  such that  $A_t^+ f$  is an integrable function on  $[0, t]$  for all  $t \geq 0$ , the process

$$f(N(t)) - f(N(0)) - \int_0^t A_s^+ f(N(s)) ds \tag{4.28}$$

is an  $\mathcal{F}_t$ -martingale. Evaluating the earlier limit gives that for each  $t \geq 0$  and  $n \in \mathbb{N} \cup \{0\}$ ,

$$A_t^+ f(n) = b(t)n(f(n+1) - f(n)) + d(t)n(f(n-1) - f(n)). \tag{4.29}$$

It follows that the process

$$\tilde{M}(t) = N(t) - N(0) - \int_0^t (b(s) - d(s))N(s) ds$$

is a right-continuous  $\mathcal{F}_t$ -martingale with  $\tilde{M}(0) = 0$ . By Doob's optional stopping

theorem,  $\mathbb{E}(\tilde{M}(t)) = 0$  for any  $t \geq 0$ , which implies

$$\mathbb{E}(N(t)) = \mathbb{E}(N(0)) + \int_0^t (b(s) - d(s))\mathbb{E}(N(s)) ds.$$

Thus  $\mathbb{E}(N(t))$  satisfies a nonhomogeneous Volterra equation of the second kind. This can be solved using Neumann series to give

$$\mathbb{E}(N(t)) = \mathbb{E}(N(0))e^{\int_0^t (b(s)-d(s)) ds} = \mathbb{E}(N(0))e^{-\rho(t)},$$

which can be shown to be the unique solution. Now using the Markov property we have for any  $0 < s < t$ :

$$\begin{aligned} \mathbb{E}(M(t)|\mathcal{F}_s) &= \mathbb{E}(N(t)e^{\rho(t)}|\mathcal{F}_s) = e^{\rho(s)}e^{-\int_s^t (b(s)-d(s)) ds} \mathbb{E}(N(t)|\mathcal{F}_s) \\ &= e^{\rho(s)}e^{-\int_s^t (b(s)-d(s)) ds} \mathbb{E}(N(s)|\mathcal{F}_s)e^{\int_s^t (b(s)-d(s)) ds} \\ &= N(s)e^{\rho(s)} = M(s), \end{aligned}$$

which shows that  $M(t)$  is a martingale. To show that  $M(t)$  is square integrable, applying any element of the set of right-generators  $(A_t^+)$  to the function  $f(n) = n^2$  and using (4.28) and (4.29) implies that the process

$$\tilde{M}_2(t) = N(t)^2 - N(0)^2 - \int_0^t (b(s) + d(s))N(s) ds + 2 \int_0^t (d(s) - b(s))N(s)^2 ds$$

is a right-continuous  $\mathcal{F}_t$ -martingale with  $\tilde{M}_2(0) = 0$ . Using Doob's optional stopping theorem as earlier,

$$\begin{aligned} \mathbb{E}(N(t)^2) &= \mathbb{E}(N(0)^2) + \int_0^t (b(s) + d(s))\mathbb{E}(N(s)) ds - 2 \int_0^t (d(s) - b(s))\mathbb{E}(N(s)^2) ds \\ &= \mathbb{E}(N(0)^2) + \mathbb{E}(N(0)) \int_0^t (b(s) + d(s))e^{-\rho(s)} ds - 2 \int_0^t (d(s) - b(s))\mathbb{E}(N(s)^2) ds. \end{aligned}$$

In view of the above representation, the function  $\varphi(t) = \mathbb{E}(N(t)^2)$  is absolutely continuous and hence  $\varphi$  satisfies the first-order linear ODE:

$$\frac{d\varphi}{dt} = \mathbb{E}(N(0))(b(t) + d(t))e^{-\rho(t)} - 2(d(t) - b(t))\varphi(t).$$

Solving the above using an integrating factor gives

$$\varphi(t) = \mathbb{E}(N(0)) \left( e^{-2\rho(t)} \int_0^t (b(s) + d(s))e^{\rho(s)} ds + e^{-2\rho(t)} \right)$$

which can be shown to be the unique solution. The second moment of  $M(t)$  is therefore given by

$$\mathbb{E}(M(t)^2) = e^{2\rho(t)}\mathbb{E}(N(t)^2) = \mathbb{E}(N(0)) \left( 1 + \int_0^t (b(s) + d(s))e^{\rho(s)} ds \right)$$

Thus by integrability of  $e^{\rho(s)}$  as well as boundedness of  $b(s)$  and  $d(s)$  we have

$$\sup_{n \geq 1} \mathbb{E}(M(t)^2) < \infty.$$

□

**Lemma 4.6.2.** *Let  $N(t)$  be a periodic birth-death process with time-dependent birth and death intensities  $b(t)$  and  $d(t)$ , such that at time  $s > 0$  there is one individual. Let  $A_s$  denote the event of eventual extinction and define the function  $\rho(s, u) = \int_s^u (d(v) - b(v)) dv$  for  $u > s$ . Then for all  $t > s$  and  $j \geq 1$  we have*

$$\mathbb{P}(N(t) = j | A_s^c) \leq \frac{\left( \int_s^t e^{-\rho(u,t)} b(u) du \right)^{j-1} \left( 1 + \int_s^\infty e^{\rho(s,u)} d(u) ds \right)}{\left( 1 + \int_s^t e^{\rho(s,u)} d(u) du \right) \left( 1 + \int_s^t e^{-\rho(u,t)} b(u) du \right)^j}.$$

*Proof.* Following [7], for times  $0 \leq s < t < \infty$ , and  $j \in \mathbb{N} \cup \{0\}$ , let  $P_j(s, t) = \mathbb{P}_{s,1}(N(t) = j) = \mathbb{P}(N(t) = j | N(s) = 1)$ . Recall that

$$\mathbb{P}_{s,1}(A_s) = \frac{\int_s^\infty e^{\rho(s,u)} d(u) du}{1 + \int_s^\infty e^{\rho(s,u)} d(u) du}$$

where for  $u > s$ ,  $\rho(s, u) = \int_s^u (d(v) - b(v)) dv$ . Then define the probability measure  $\mathbb{Q}_{s,1}$  via

$$\mathbb{Q}_{s,1}(N(t) = j) = \mathbb{P}_{s,1}(N(t) = j | A_s^c) = \frac{\mathbb{P}_{s,1}(N(t) = j, A_s^c)}{\mathbb{P}_{s,1}(A_s^c)}$$

and set  $\hat{P}_j(s, t) = \mathbb{Q}_{s,1}(N(t) = j)$  for  $j \in \mathbb{N} \cup \{0\}$ . Then we have for  $j \geq 1$ ,

$$\hat{P}_j(s, t) = \frac{\mathbb{P}_{s,1}(N(t) = j, A_s^c)}{\mathbb{P}_{s,1}(A_s^c)} \leq \frac{\mathbb{P}_{s,1}(N(t) = j)}{1 - \mathbb{P}_{s,1}(A_s)} = P_j(s, t) \left( 1 + \int_s^\infty e^{\rho(s,u)} d(u) ds \right)$$

and  $\hat{P}_0(s, t) = 0$ . From equations (5) and (6) in [7] we know that  $P_j(s, t) = (1 -$

$P_0(s, t)(1 - \eta(s, t))\eta(s, t)^{j-1}$ , where  $\eta(s, t)$  is the function

$$\eta(s, t) = \frac{\int_s^t e^{-\rho(u, t)} b(u) du}{1 + \int_s^t e^{-\rho(u, t)} b(u) du}$$

and from [24] we have that

$$P_0(s, t) = \frac{\int_s^t e^{\rho(s, u)} d(u) du}{1 + \int_s^t e^{\rho(s, u)} d(u) du}.$$

Therefore,

$$\begin{aligned} \hat{P}_j(s, t) &= (1 - P_0(s, t))(1 - \eta(s, t))\eta(s, t)^{j-1} \left(1 + \int_s^\infty e^{\rho(s, u)} d(u) ds\right) \\ &\leq \frac{\left(\int_s^t e^{-\rho(u, t)} b(u) du\right)^{j-1} \left(1 + \int_s^\infty e^{\rho(s, u)} d(u) ds\right)}{\left(1 + \int_s^t e^{\rho(s, u)} d(u) du\right) \left(1 + \int_s^t e^{-\rho(u, t)} b(u) du\right)^j} \end{aligned} \quad (4.30)$$

□

**Lemma 4.6.3.** *Suppose  $f(s) \geq 0$  and  $f'(s) \geq 0$  for all  $s \geq 0$ , and that  $f$  is locally Lipschitz continuous. Then for any  $T > 0$  and  $0 \leq V < V + 2U \leq T$  we have*

$$\int_{V+U}^{V+2U} f(s) ds - \int_V^{V+U} f(s) ds \leq QU^2$$

for some  $Q > 0$ .

*Proof.* Without loss of generality, assume  $V = 0$ . Then

$$\begin{aligned} \int_U^{2U} f(s) ds - \int_0^U f(s) ds &= \int_U^{2U} f(s) ds - Uf(U) - \left(\int_0^U f(s) ds - Uf(U)\right) \\ &= \int_U^{2U} (f(s) - f(U)) ds + \int_0^U (f(U) - f(s)) ds \\ &\leq Q \int_U^{2U} (s - U) ds + Q \int_0^U (U - s) ds \\ &= Q \left(\frac{(2U - U)^2}{2} - \frac{(U - U)^2}{2}\right) + Q \left(-\frac{(U - U)^2}{2} + \frac{(0 - U)^2}{2}\right) \\ &= QU^2 \end{aligned}$$

for some  $Q > 0$ .

□

## Chapter 5

# Invasion with Delay

We now no longer assume a periodicity in the birth and death rates but instead look at the situation where one of two competing mutants is introduced into a nutrient-rich environment at some time after the introduction of the first mutant. Where one mutant is already resident, i.e. where at time zero it has population density equal to its steady state value from equation (1.10), this question has already been studied in the papers of Champagnat [8] and Billiard and Smadi [5]. In particular, Lemma 3.3 in [5] implies the third bullet point in our main theorem (Theorem 5.1.2), however they do not prove this lemma. The question we are interested in is whether despite a negative invasion fitness, the type 1 population can become resident if neither population is initially resident. It is clear that for this to occur, for some  $\varepsilon > 0$  the type 1 population must reach size  $\varepsilon K$  (which we refer to as establishment) before the type 0 population does. If  $b_1 - d_1 > b_0 - d_0$  then it is also clear that assuming both populations begin growing at the same time, the type 1 population will reach size  $\varepsilon K$  first. However if we introduce a delay in the introduction of the type 1 individuals (the fitter mutants), the answer becomes uncertain. In addition, to extend the results of [5], we consider how starting populations at sizes of order  $K^\gamma$  for  $\gamma \in (0, 1)$  affects the dynamics. This also draws a closer link to the second experiment described in the introduction, in which A289C cells were introduced into an existing plate of A683C cells but at a starting size on the order of  $10^7$  cells. To investigate this question, our model setup is as follows:

Consider a pair of coupled logistic birth-death processes  $(N_0(t), N_1(t))_{t \geq 0}$  such that at time  $t \in [0, \infty)$ ,

- Each type 0 individual gives birth at rate  $b_0$  and dies at rate  $d_0 + \frac{C_{0,0}N_0(t)}{K} + \frac{C_{0,1}N_1(t)}{K}$ .

- Each type 1 individual gives birth at rate  $b_1$  and dies at rate  $d_1 + \frac{C_{1,1}N_1(t)}{K} + \frac{C_{0,1}(t)N_0(t)}{K}$

where  $K > 0$  is an integer parameter representing the system size. Now suppose that  $N_0(0) = K^{\gamma_0}$  for some  $\gamma_0 \geq 0$  and  $N_1(0) = 0$ . We also let  $T_D$  denote the time delay in the introduction of the type 1 individuals so that  $N_1(t) = 0$  for all  $t < T_D$  and specify that  $N_1(T_D) = K^{\gamma_1}$  for some  $\gamma_1 \geq 0$ . For a fixed  $\alpha > 0$ , we choose the delay deterministically as  $T_D = \alpha \log K$ . The question is how large must  $\alpha$  be to ensure the type 1 population is unable to invade the type 0 population despite the type 1 population being fitter, i.e.  $0 < b_0 - d_0 < b_1 - d_1$ .

## 5.1 Probability of Invasion

For each type  $i \in \{0, 1\}$  and  $R > 0$ , define the stopping times

$$T_R^{(i)} = \inf \{t \geq 0 : N_i(t) \geq R\}$$

and

$$T_0^{(i)} = \inf \{t \geq 0 : N_i(t) = 0\}.$$

Then we can introduce a birth-death process  $N_0^+(t)$  with birth rate  $b_0$  and death rate  $d_0^+ = d_0$ , as well as a birth-death process  $N_0^-(t)$  with birth rate  $b_0$  and death rate  $d_0^- = d_0 + (C_{0,0} + C_{0,1})\varepsilon$  for some  $\varepsilon > 0$ , coupled with  $N_0(t)$  such that

$$\mathbb{P}\left(N_0^-(t) \leq N_0(t) \leq N_0^+(t) \text{ for all } t \in [0, T_{\varepsilon K}^{(0)} \wedge T_{\varepsilon K}^{(1)}]\right) = 1.$$

Similarly, we can introduce a birth-death process  $N_1^+(t)$  with birth rate  $b_1$  and death rate  $d_1^+ = d_1$ , as well as a birth-death process  $N_1^-(t)$  with birth rate  $b_1$  and death rate  $d_1^- = d_1 + (C_{0,1} + C_{1,1})\varepsilon$ , coupled with  $N_1(t)$  such that

$$\mathbb{P}\left(N_1^-(t) \leq N_1(t) \leq N_1^+(t) \text{ for all } t \in [T_{\varepsilon K}^{(0)} \wedge T_D, T_{\varepsilon K}^{(0)} \wedge T_{\varepsilon K}^{(1)}]\right) = 1.$$

As for the original processes, for  $R > 0$ ,  $i \in \{0, 1\}$  and  $* \in \{-, +\}$  we define the stopping times

$$T_R^{(i,*)} = \inf \{t \geq 0 : N_i^*(t) \geq R\}$$

and

$$T_0^{(i,*)} = \inf \{t \geq 0 : N_i^*(t) = 0\}.$$

We now introduce a preliminary lemma that will enable us to estimate the time taken for a standard birth-death process to reach level  $\varepsilon K$  given that it is started at level  $K^\gamma$ ,  $\gamma > 0$ .

**Lemma 5.1.1.** *Let  $N(t)$  denote a standard birth-death process of birth rate  $b$  and death rate  $d < b$  with initial condition  $N(0) = K^\gamma$  for some  $\gamma \in (0, 1)$ , and let  $\varepsilon \in (0, 1]$ . Then on the event of non-extinction, for any  $\delta > 0$ , the time taken to reach level  $\varepsilon K$ , denoted  $T_{\varepsilon K}$ , satisfies*

$$\mathbb{P}\left(\left|\frac{T_{\varepsilon K}}{\log(\varepsilon K)} - \frac{1-\gamma}{b-d}\right| > \delta \mid T_0 = \infty\right) \rightarrow 0$$

as  $K \rightarrow \infty$ .

*Proof.* Define the event of non-extinction  $B_K = \{T_0 = \infty\}$  and suppose we order the  $K^\gamma$  starting individuals from left to right. Consider the left-most  $\varepsilon^\gamma K^\gamma$  individuals and let  $T'_0$  denote the extinction time of this subpopulation. We let  $C_K = \{T'_0 = \infty\}$  denote the event that this subpopulation does not become extinct, i.e. at least one of these left-most individuals has a lineage (sequence of descendants) of infinite length. Note that  $C_K \subseteq B_K$ . Also, let  $T'_{\varepsilon K}$  denote the time taken for the number of descendants of the left-most subpopulation of  $\varepsilon^\gamma K^\gamma$  starting individuals to reach level  $\varepsilon K$ , and note that on the event  $C_K$  we have  $T_{\varepsilon K} \leq T'_{\varepsilon K}$  almost surely. Denote the event of interest by

$$E_K = \left\{ \left| \frac{T_{\varepsilon K}}{\log(\varepsilon K)} - \frac{1-\gamma}{b-d} \right| > \delta \right\}$$

which we decompose into disjoint sets  $E_K = E_K^1 \cup E_K^2$  where

$$E_K^1 = \left\{ \frac{T_{\varepsilon K}}{\log(\varepsilon K)} < \frac{1-\gamma}{b-d} - \delta \right\} \quad \text{and} \quad E_K^2 = \left\{ \frac{T_{\varepsilon K}}{\log(\varepsilon K)} > \frac{1-\gamma}{b-d} + \delta \right\}.$$

First note that since  $C_K \subseteq B_K$ ,

$$\mathbb{P}(E_K^2 \mid B_K) \leq \mathbb{P}(E_K^2 \mid C_K) \left( \frac{\mathbb{P}(C_K)}{\mathbb{P}(B_K)} \right) + \frac{\mathbb{P}(C_K^c)}{\mathbb{P}(B_K)}.$$

We have

$$\begin{aligned} \mathbb{P}(E_K^2 \mid C_K) &= \mathbb{P}\left(\frac{T_{\varepsilon K}}{\log(\varepsilon K)} > \frac{1-\gamma}{b-d} + \delta \mid C_K\right) \leq \mathbb{P}\left(\frac{T'_{\varepsilon K}}{\log(\varepsilon K)} > \frac{1-\gamma}{b-d} + \delta \mid C_K\right) \\ &\leq \mathbb{P}\left(\left|\frac{T'_{\varepsilon K}}{\log(\varepsilon K)} - \frac{1-\gamma}{b-d}\right| > \delta \mid T'_0 = \infty\right) \\ &\rightarrow 0 \end{aligned}$$



as  $K \rightarrow \infty$ , where we use Corollary 4.1.3 for the convergence in the last line. Additionally,  $\mathbb{P}(B_K) = 1 - (d/b)^{K^\gamma}$  and  $\mathbb{P}(C_K) = 1 - (d/b)^{(\varepsilon K)^\gamma}$  so as  $b > d$  we have  $\mathbb{P}(C_K)/\mathbb{P}(B_K) \rightarrow 1$  and  $\mathbb{P}(C_K^c)/\mathbb{P}(B_K) \rightarrow 0$  as  $K \rightarrow \infty$ . Therefore

$$\mathbb{P}(E_K^2 | B_K) \rightarrow 0$$

as  $K \rightarrow \infty$ . To find the limit of  $\mathbb{P}(E_K^1 | B_K)$ , choose  $\beta \in (\gamma, 1)$  such that  $\beta > 1 - \frac{\delta}{2}(b-d)$  and define  $T_{K^\beta}$  to be the first time the population reaches level  $K^\beta$ . Note that for large enough  $K$ ,  $T_{K^\beta} \leq T_{\varepsilon K}$  almost surely. Then by our choice of  $\beta$  we have

$$\begin{aligned} \mathbb{P}(E_K^1 | B_K) &= \mathbb{P}\left(\frac{T_{\varepsilon K}}{\log(\varepsilon K)} < \frac{1-\gamma}{b-d} - \delta \mid T_0 = \infty\right) \leq \mathbb{P}\left(\frac{T_{K^\beta}}{\log K} < \frac{1-\gamma}{b-d} - \delta \mid T_0 = \infty\right) \\ &\leq \mathbb{P}\left(\frac{T_{K^\beta}}{\log K} < \frac{1-\gamma}{b-d} - \delta \mid T_0 = \infty\right) \leq \mathbb{P}\left(\frac{T_{K^\beta}}{\log K} < \frac{\beta-\gamma}{b-d} - \frac{\delta}{2} \mid T_0 = \infty\right) \\ &\leq \mathbb{P}\left(\left|\frac{T_{K^\beta}}{\log K} - \frac{\beta-\gamma}{b-d}\right| > \frac{\delta}{2} \mid T_0 = \infty\right) \rightarrow 0 \end{aligned}$$

where we again use Corollary 4.1.3 for the convergence in the last line. Hence,

$$\mathbb{P}(E_K | B_K) = \mathbb{P}(E_K^1 | B_K) + \mathbb{P}(E_K^2 | B_K) \rightarrow 0$$

as  $K \rightarrow \infty$ . □

Since the eventual fixation of one of the types after establishment will depend on the stable steady state(s) of a deterministic ODE, we now focus on the probability of establishment. Below we restate Theorem 1.2.2 from the introduction and note that we require  $\gamma_1$  to be sufficiently large with respect to  $\gamma_0$ , as specified by condition (5.1), otherwise the type 1 will fail to invade for all  $\alpha > 0$ . Note that this condition is trivially satisfied if  $\gamma_0 < 1 - \frac{b_0-d_0}{b_1-d_1}$ , in particular it is satisfied in the case where the type 0 population starts with 1 individual.

**Theorem 5.1.2.** *Suppose  $0 < b_0 - d_0 < b_1 - d_1$  and let  $\gamma_0, \gamma_1 \in [0, 1)$  be such that*

$$(\gamma_1 - 1) > (\gamma_0 - 1) \left( \frac{b_1 - d_1}{b_0 - d_0} \right). \quad (5.1)$$

*If  $0 < \alpha < \frac{1-\gamma_0}{b_0-d_0} - \frac{1-\gamma_1}{b_1-d_1}$  there exists  $R > 0$  such that for  $\varepsilon > 0$  sufficiently small, we have as  $K \rightarrow \infty$ ,*

- if  $\gamma_0 \geq 0, \gamma_1 > 0$ ,

$$\mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}) \rightarrow 1 \quad \text{and} \quad \mathbb{P}(T_{\varepsilon K}^{(0)} < T_{\varepsilon K}^{(1)}) \rightarrow 0$$

- if  $\gamma_0 > 0, \gamma_1 = 0$ ,

$$\left| \mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}) - \left(1 - \frac{d_1}{b_1}\right) \right| < R\varepsilon \quad \text{and} \quad \left| \mathbb{P}(T_{\varepsilon K}^{(0)} < T_{\varepsilon K}^{(1)}) - \frac{d_1}{b_1} \right| < R\varepsilon$$

- if  $\gamma_0 = 0, \gamma_1 = 0$ ,

$$\left| \mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}) - \left(1 - \frac{d_1}{b_1}\right) \right| < R\varepsilon \quad , \quad \left| \mathbb{P}(T_{\varepsilon K}^{(0)} < T_{\varepsilon K}^{(1)}) - \left(1 - \frac{d_0}{b_0}\right) \frac{d_1}{b_1} \right| < R\varepsilon$$

$$\text{and} \quad \left| \mathbb{P}(T_0^{(0)} \vee T_0^{(1)} < T_{\varepsilon K}^{(0)} \wedge T_{\varepsilon K}^{(1)}) - \frac{d_0 d_1}{b_0 b_1} \right| < R\varepsilon$$

If instead  $\alpha > \frac{1-\gamma_0}{b_0-d_0} - \frac{1-\gamma_1}{b_1-d_1}$  there exists  $R > 0$  such that for  $\varepsilon > 0$  sufficiently small we have as  $K \rightarrow \infty$ ,

- if  $\gamma_0 > 0, \gamma_1 \geq 0$ ,

$$\mathbb{P}(T_{\varepsilon K}^{(0)} < T_{\varepsilon K}^{(1)}) \rightarrow 1 \quad \text{and} \quad \mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}) \rightarrow 0$$

- if  $\gamma_0 = 0, \gamma_1 > 0$

$$\left| \mathbb{P}(T_{\varepsilon K}^{(0)} < T_{\varepsilon K}^{(1)}) - \left(1 - \frac{d_0}{b_0}\right) \right| < R\varepsilon \quad \text{and} \quad \left| \mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}) - \frac{d_0}{b_0} \right| < R\varepsilon$$

- if  $\gamma_0 = 0, \gamma_1 = 0$ ,

$$\left| \mathbb{P}(T_{\varepsilon K}^{(0)} < T_{\varepsilon K}^{(1)}) - \left(1 - \frac{d_0}{b_0}\right) \right| < R\varepsilon \quad , \quad \left| \mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}) - \frac{d_0}{b_0} \left(1 - \frac{d_1}{b_1}\right) \right| < R\varepsilon$$

$$\text{and} \quad \left| \mathbb{P}(T_0^{(0)} \vee T_0^{(1)} < T_{\varepsilon K}^{(0)} \wedge T_{\varepsilon K}^{(1)}) - \frac{d_0 d_1}{b_0 b_1} \right| < R\varepsilon$$

In particular, if  $\alpha > \frac{1-\gamma_0}{b_0-d_0}$  and  $\gamma_0 > 0$ ,

$$\mathbb{P}(T_{\varepsilon K}^{(0)} < T_D) \rightarrow 1$$

as  $K \rightarrow \infty$ .

*Proof.* We proceed initially in a similar way for each of the cases; in the course of the proof it will be shown how each of the cases leads to different limits. The first task is to estimate the size of the type 0 population at time  $T_D$ . To do this, we can split over the event that the process  $N_0^+(t)$  eventually goes extinct. We have

$$\mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}) = \mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}, T_0^{(0,+)} < \infty) + \mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}, T_0^{(0,+)} = \infty) \quad (5.2)$$

and similarly using the process  $N_0^-(t)$ ,

$$\mathbb{P}(T_{\varepsilon K}^{(0)} < T_{\varepsilon K}^{(1)}) = \mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}, T_0^{(0,-)} < \infty) + \mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}, T_0^{(0,-)} = \infty). \quad (5.3)$$

(i)  $\gamma_0 = 0$ ,  $\mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)})$

If  $\gamma_0 = 0$  we need to consider the first of the two probabilities on the right hand side of equation (5.2). First note that the couplings  $N_0(t) \leq N_0^+(t)$  and  $N_1(t) \geq N_1^-(t)$  hold almost surely for all times  $t \leq T_{\varepsilon K}^{(1)} \wedge T_{\varepsilon K}^{(0)}$ . Therefore, if  $T_{\varepsilon K}^{(1,-)} < T_{\varepsilon K}^{(0,+)}$ , it follows almost surely that  $T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}$ . Indeed, if it were not the case and  $T_{\varepsilon K}^{(0)} \leq T_{\varepsilon K}^{(1)}$ , then by the coupling for the type 0 population, we would have  $T_{\varepsilon K}^{(0,+)} \leq T_{\varepsilon K}^{(0)}$ . But by assumption  $T_{\varepsilon K}^{(1,-)} < T_{\varepsilon K}^{(0,+)}$ , so the coupling for the type 1 population gives  $T_{\varepsilon K}^{(1)} \leq T_{\varepsilon K}^{(1,-)}$  yielding a contradiction. We will use this argument implicitly throughout the proof. Note that

$$\begin{aligned} & \mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}, T_0^{(0,+)} < \infty) \geq \mathbb{P}(T_{\varepsilon K}^{(1,-)} < T_{\varepsilon K}^{(0,+)}, T_0^{(0,+)} < \infty) \\ & \geq \mathbb{P}(T_{\varepsilon K}^{(1,-)} < T_{\varepsilon K}^{(0,+)}, T_0^{(1,-)} = \infty, T_0^{(0,+)} < \infty) \\ & \geq \mathbb{P}(T_{\varepsilon K}^{(1,-)} < T_{\varepsilon K}^{(0,+)}, T_0^{(1,-)} = \infty, T_0^{(0,+)} < \infty, T_0^{(0,+)} < T_{\varepsilon K}^{(0,+)}) \\ & = \mathbb{P}(T_{\varepsilon K}^{(1,-)} < \infty, T_0^{(1,-)} = \infty, T_0^{(0,+)} < \infty, T_0^{(0,+)} < T_{\varepsilon K}^{(0,+)}) \quad (5.4) \\ & = \mathbb{P}(T_{\varepsilon K}^{(1,-)} < \infty \mid T_0^{(1,-)} = \infty) \mathbb{P}(T_0^{(1,-)} = \infty) \\ & \times \mathbb{P}(T_0^{(0,+)} < T_{\varepsilon K}^{(0,+)} \mid T_0^{(0,+)} < \infty) \mathbb{P}(T_0^{(0,+)} < \infty) \\ & = \mathbb{P}(T_0^{(0,+)} < T_{\varepsilon K}^{(0,+)} \mid T_0^{(0,+)} < \infty) \mathbb{P}(T_0^{(0,+)} < \infty) \mathbb{P}(T_0^{(1,-)} = \infty) \end{aligned}$$

Note that by the formula for the extinction probability of a population started with

$K^{\gamma_1}$  individuals, if  $\gamma_1 > 0$ ,

$$\mathbb{P}(T_0^{(1,-)} = \infty) = 1 - \left( \frac{d_1 + \varepsilon(C_{0,1} + C_{1,1})}{b_1} \right)^{K^{\gamma_1}} \rightarrow 1$$

as  $K \rightarrow \infty$  and  $\mathbb{P}(T_0^{(1,-)} = \infty) = 1 - (d_1 + \varepsilon(C_{0,1} + C_{1,1})) / b_1$  if  $\gamma_1 = 0$ . Additionally, since  $\gamma_0 = 0$  we have  $\mathbb{P}(T_0^{(0,+)} < \infty) = d_0 / b_0$ . Also, a supercritical birth-death process with birth rate  $b_0$  and death rate  $d_0$  conditioned on extinction has sample paths distributed according to a subcritical birth-death process of birth rate  $d_0$  and death rate  $b_0$  (see Theorem 3 of Chapter 1, Section 12 in [2]). Using the formula for the probability of hitting level 0 before reaching level  $\varepsilon K$  when starting at level 1 (see part 1 of Lemma 4.1.1), we have

$$\mathbb{P}(T_0^{(0,+)} < T_{\varepsilon K}^{(0,+)} \mid T_0^{(0,+)} < \infty) = 1 - \left( \frac{\frac{b_0}{d_0} - 1}{\left(\frac{b_0}{d_0}\right)^{\varepsilon K} - 1} \right) \rightarrow 1$$

as  $K \rightarrow \infty$ . Having computed the limits for each of the three terms in the product in (5.4), we get the lower bounds

$$\begin{aligned} \liminf_{K \rightarrow \infty} \mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}, T_0^{(0,+)} < \infty) &\geq \frac{d_0}{b_0} \quad \text{if } \gamma_1 > 0 \\ \liminf_{K \rightarrow \infty} \mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}, T_0^{(0,+)} < \infty) &\geq \frac{d_0}{b_0} \left( 1 - \frac{d_1}{b_1} - \frac{\varepsilon C_{1,1}}{b_1} \right) \quad \text{if } \gamma_1 = 0. \end{aligned} \quad (5.5)$$

Next we need a lower bound for the second term in (5.2),  $\mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}, T_0^{(0,+)} = \infty)$ . Using the couplings as before, we have that

$$\begin{aligned} \mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}, T_0^{(0,+)} = \infty) &\geq \mathbb{P}(T_{\varepsilon K}^{(1,-)} < T_{\varepsilon K}^{(0,+)}, T_0^{(0,+)} = \infty) \\ &= \mathbb{P}(T_{\varepsilon K}^{(1,-)} < T_{\varepsilon K}^{(0,+)} \mid T_0^{(0,+)} = \infty) \mathbb{P}(T_0^{(0,+)} = \infty) \end{aligned}$$

Given  $\alpha < \frac{1-\gamma_0}{b_0-d_0} - \frac{1-\gamma_1}{b_1-d_1}$ , for  $\gamma_0 = 0$  we choose  $\varepsilon > 0$  small enough such that there exists  $\delta > 0$  such that  $\alpha + \frac{1-\gamma_1}{b_1-d_1-(C_{0,1}+C_{1,1})\varepsilon} + \delta < \frac{1-\gamma_0}{b_0-d_0} - \delta$ . Note that given  $\varepsilon \leq 1$ , for any  $\delta > 0$  we can choose  $K$  large enough so that  $\frac{\log K}{\log(\varepsilon K)} = \frac{\log K}{\log K - \log(1/\varepsilon)} \leq 1 + \frac{\delta}{2\alpha}$ .

We therefore have using independence of the processes  $N_1^-(t)$  and  $N_0^+(t)$ ,

$$\begin{aligned}
& \mathbb{P}(T_{\varepsilon K}^{(1,-)} < T_{\varepsilon K}^{(0,+)} \mid T_0^{(0,+)} = \infty) \geq \mathbb{P}\left(\frac{T_{\varepsilon K}^{(1,-)}}{\log(\varepsilon K)} < \frac{T_{\varepsilon K}^{(0,+)}}{\log(\varepsilon K)} \mid T_0^{(0,+)} = \infty\right) \\
& \geq \mathbb{P}\left(\frac{T_{\varepsilon K}^{(1,-)}}{\log(\varepsilon K)} \leq \left(\alpha + \frac{1-\gamma_1}{b_1-d_1-(C_{0,1}+C_{1,1})\varepsilon} + \delta\right) < \frac{1-\gamma_0}{b_0-d_0} - \delta \leq \frac{T_{\varepsilon K}^{(0,+)}}{\log(\varepsilon K)} \mid T_0^{(0,+)} = \infty\right) \\
& \geq \mathbb{P}\left(\frac{T_{\varepsilon K}^{(1,-)}}{\log(\varepsilon K)} \leq \left(\alpha + \frac{1-\gamma_1}{b_1-d_1-(C_{0,1}+C_{1,1})\varepsilon} + \delta\right) < \frac{1-\gamma_0}{b_0-d_0} - \delta \leq \frac{T_{\varepsilon K}^{(0,+)}}{\log(\varepsilon K)} \mid T_0^{(1-)} = \infty, T_0^{(0,+)} = \infty\right) \\
& \times \mathbb{P}(T_0^{(1-)} = \infty, T_0^{(0,+)} = \infty \mid T_0^{(0,+)} = \infty) \\
& \geq \mathbb{P}\left(\frac{T_{\varepsilon K}^{(1,-)}}{\log(\varepsilon K)} \leq \alpha + \frac{1-\gamma_0}{b_1-d_1-(C_{0,1}+C_{1,1})\varepsilon} + \delta \mid T_0^{(1-)} = \infty\right) \\
& \times \mathbb{P}\left(\frac{T_{\varepsilon K}^{(0,+)}}{\log(\varepsilon K)} \geq \frac{1-\gamma_0}{b_0-d_0} - \delta \mid T_0^{(0,+)} = \infty\right) \mathbb{P}(T_0^{(1-)} = \infty, T_0^{(0,+)} = \infty \mid T_0^{(0,+)} = \infty) \\
& \geq \mathbb{P}\left(\frac{T_{\varepsilon K}^{(1,-)}}{\log(\varepsilon K)} \leq \alpha \left(\frac{\log K}{\log(\varepsilon K)}\right) + \frac{1-\gamma_1}{b_1-d_1-(C_{0,1}+C_{1,1})\varepsilon} + \frac{\delta}{2} \mid T_0^{(1-)} = \infty\right) \\
& \times \mathbb{P}\left(\frac{T_{\varepsilon K}^{(0,+)}}{\log(\varepsilon K)} \geq \frac{1-\gamma_0}{b_0-d_0} - \delta \mid T_0^{(0,+)} = \infty\right) \\
& \times \mathbb{P}(T_0^{(1-)} = \infty, T_0^{(0,+)} = \infty \mid T_0^{(0,+)} = \infty)
\end{aligned}$$

Note that in both cases ( $\gamma_1 \geq 0$ ) we have by Lemma 5.1.1,

$$\begin{aligned}
& \mathbb{P}\left(\frac{T_{\varepsilon K}^{(1,-)}}{\log(\varepsilon K)} \leq \alpha \left(\frac{\log K}{\log(\varepsilon K)}\right) + \frac{1-\gamma_1}{b_1-d_1-(C_{0,1}+C_{1,1})\varepsilon} + \frac{\delta}{2} \mid T_0^{(1-)} = \infty\right) \\
& = \mathbb{P}\left(\frac{T_{\varepsilon K}^{(1,-)}}{\log(\varepsilon K)} \leq \frac{T_D}{\log(\varepsilon K)} + \frac{1-\gamma_1}{b_1-d_1-(C_{0,1}+C_{1,1})\varepsilon} + \frac{\delta}{2} \mid T_0^{(1-)} = \infty\right) \rightarrow 1
\end{aligned}$$

as  $K \rightarrow \infty$ , and since  $\gamma_0 = 0$  we have from part 3 of Lemma A.1 in [5],

$$\mathbb{P}\left(\frac{T_{\varepsilon K}^{(0,+)}}{\log(\varepsilon K)} \geq \frac{1-\gamma_0}{b_0-d_0} - \delta \mid T_0^{(0,+)} = \infty\right) \rightarrow 1$$

as  $K \rightarrow \infty$ . Again by independence of  $N_1^-(t)$  and  $N_0^+(t)$ , we have

$$\begin{aligned}
& \mathbb{P}(T_0^{(1,-)} = \infty, T_0^{(0,+)} = \infty \mid T_0^{(0,+)} = \infty) \\
& = \mathbb{P}(T_0^{(1,-)} = \infty) \mathbb{P}(T_0^{(0,+)} = \infty \mid T_0^{(0,+)} = \infty) \\
& = \left(1 - \left(\frac{d_1 + \varepsilon(C_{0,1} + C_{1,1})}{b_1}\right)^{K^{\gamma_1}}\right),
\end{aligned}$$

which clearly converges to 1 if  $\gamma_1 > 0$  and converges to  $1 - (d_1/b_1) - \varepsilon(C_{0,1} + C_{1,1})/d_1$

if  $\gamma_1 = 0$ . Together with the fact that due to  $\gamma_0 = 0$ ,  $\mathbb{P}(T_0^{(0,+)} = \infty) = 1 - d_0/b_0$ , we get that

$$\begin{aligned} \liminf_{K \rightarrow \infty} \mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}, T_0^{(0,+)} = \infty) &\geq 1 - \frac{d_0}{b_0} \quad \text{if } \gamma_1 > 0, \text{ and} \\ \liminf_{K \rightarrow \infty} \mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}, T_0^{(0,+)} = \infty) &= \left(1 - \frac{d_0}{b_0}\right) \left(1 - \frac{d_1}{b_1} - \frac{\varepsilon(C_{0,1} + C_{1,1})}{b_1}\right) \\ &\text{if } \gamma_1 = 0. \end{aligned} \quad (5.6)$$

Pulling together (5.5) and (5.6), and referring to equation (5.2) gives us that

$$\mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}) \rightarrow 1 \quad \text{if } \gamma_0 = 0, \gamma_1 > 0$$

and

$$\liminf_{K \rightarrow \infty} \mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}) \geq \left(1 - \frac{d_1}{b_1}\right) - R\varepsilon \quad \text{if } \gamma_0 = 0, \gamma_1 = 0$$

for some  $R > 0$ .

(ii)  $\gamma_0 = 0$ ,  $\mathbb{P}(T_{\varepsilon K}^{(0)} < T_{\varepsilon K}^{(1)})$

In view of the previous section, if  $\gamma_1 > 0$  we automatically have

$$\mathbb{P}(T_{\varepsilon K}^{(0)} < T_{\varepsilon K}^{(1)}) \rightarrow 0$$

so we only need to concentrate on the case  $\gamma_1 = 0$ . Unlike in the previous section, we can neglect the first term in (5.3) and concentrate on the second term (this will be sufficient for the lower bound we need). Note that by independence of the processes  $N_0^-(t)$  and  $N_1^+(t)$  we have,

$$\begin{aligned} \mathbb{P}(T_{\varepsilon K}^{(0)} < T_{\varepsilon K}^{(1)}, T_0^{(0,-)} = \infty) &\geq \mathbb{P}(T_{\varepsilon K}^{(0,-)} < T_{\varepsilon K}^{(1,+)}, T_0^{(0,-)} = \infty) \\ &\geq \mathbb{P}(T_{\varepsilon K}^{(0,-)} < T_{\varepsilon K}^{(1,+)}, T_0^{(0,-)} = \infty, T_0^{(1,+)} < T_{\varepsilon K}^{(1,+)}) \\ &= \mathbb{P}(T_{\varepsilon K}^{(0,-)} < \infty, T_0^{(0,-)} = \infty, T_0^{(1,+)} < T_{\varepsilon K}^{(1,+)} \\ &= \mathbb{P}(T_{\varepsilon K}^{(0,-)} < \infty \mid T_0^{(0,-)} = \infty) \mathbb{P}(T_0^{(0,-)} = \infty) \mathbb{P}(T_0^{(1,+)} < T_{\varepsilon K}^{(1,+)} \\ &= \mathbb{P}(T_0^{(0,-)} = \infty) \mathbb{P}(T_0^{(1,+)} < T_{\varepsilon K}^{(1,+)}). \end{aligned} \quad (5.7)$$

By the formula for the extinction time of a birth death process starting from  $K^{\gamma_0} = 1$  individual (see part 2 of Lemma A.1 in [5]) we have

$$\mathbb{P}(T_0^{(0,-)} = \infty) = 1 - \frac{d_0 + \varepsilon(C_{0,0} + C_{0,1})}{b_0}. \quad (5.8)$$

By part 2 of Lemma A.1 in [5] we also have

$$\mathbb{P}(T_0^{(1,+)} < T_{\varepsilon K}^{(1,+)}) = \left(1 - \frac{1 - \frac{d_1}{b_1}}{1 - \left(\frac{d_1}{b_1}\right)^{\varepsilon K}}\right) \rightarrow \frac{d_1}{b_1} \quad (5.9)$$

as  $K \rightarrow \infty$ . Combining (5.8) and (5.9) with (5.7) we get

$$\begin{aligned} \mathbb{P}(T_{\varepsilon K}^{(0)} < T_{\varepsilon K}^{(1)}, T_0^{(0,-)} = \infty) &\rightarrow 0 \quad \text{if } \gamma_1 > 0 \\ \liminf_{K \rightarrow \infty} \mathbb{P}(T_{\varepsilon K}^{(0)} < T_{\varepsilon K}^{(1)}, T_0^{(0,-)} = \infty) &= \left(1 - \frac{d_0 + \varepsilon C_{0,0}}{b_0}\right) \frac{d_1}{b_1} \quad \text{if } \gamma_1 = 0. \end{aligned} \quad (5.10)$$

Using (5.3) we then have

$$\begin{aligned} \mathbb{P}(T_{\varepsilon K}^{(0)} < T_{\varepsilon K}^{(1)}) &\rightarrow 0 \quad \text{if } \gamma_1 > 0 \\ \liminf_{K \rightarrow \infty} \mathbb{P}(T_{\varepsilon K}^{(0)} < T_{\varepsilon K}^{(1)}) &\geq \left(1 - \frac{d_0 + \varepsilon C_{0,0}}{b_0}\right) \frac{d_1}{b_1} \quad \text{if } \gamma_1 = 0. \end{aligned} \quad (5.11)$$

**(iii)**  $\gamma_0 > 0$ ,  $\mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)})$

With  $\gamma_0 > 0$ , we have using independence of the processes  $N_0^+(t)$  and  $N_1^-(t)$ ,

$$\begin{aligned} \mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}) &\geq \mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}, T_0^{(0,+)} = \infty, T_0^{(1,-)} = \infty) \\ &\geq \mathbb{P}(T_{\varepsilon K}^{(1,-)} < T_{\varepsilon K}^{(0,+)}, T_0^{(0,+)} = \infty, T_0^{(1,-)} = \infty) \\ &= \mathbb{P}(T_{\varepsilon K}^{(1,-)} < T_{\varepsilon K}^{(0,+)} \mid T_0^{(0,+)} = \infty, T_0^{(1,-)} = \infty) \mathbb{P}(T_0^{(0,+)} = \infty) \mathbb{P}(T_0^{(1,-)} = \infty). \end{aligned} \quad (5.12)$$

As in part (i), given our condition  $\alpha < \frac{1-\gamma_0}{b_0-d_0} - \frac{1-\gamma_1}{b_1-d_1}$ , we choose  $\varepsilon > 0$  small enough such that there exists  $\delta > 0$  such that  $\alpha + \frac{1-\gamma_1}{b_1-d_1-(C_{0,1}+C_{1,1})\varepsilon} + \delta < \frac{1-\gamma_0}{b_0-d_0} - \delta$ . Note that given  $\varepsilon \leq 1$ , for any  $\delta > 0$  we can choose  $K$  large enough so that  $\frac{\log K}{\log(\varepsilon K)} = \frac{\log K}{\log K - \log(1/\varepsilon)} \leq$

$1 + \frac{\delta}{2\alpha}$ . We therefore have using independence of the processes  $N_1^-(t)$  and  $N_0^+(t)$ ,

$$\begin{aligned}
& \mathbb{P}(T_{\varepsilon K}^{(1,-)} < T_{\varepsilon K}^{(0,+)} \mid T_0^{(0,+)} = \infty, T_0^{(1,-)} = \infty) \\
& \geq \mathbb{P}\left(\frac{T_{\varepsilon K}^{(1,-)}}{\log(\varepsilon K)} \leq \left(\alpha + \frac{1-\gamma_1}{b_1-d_1-(C_{0,1}+C_{1,1})\varepsilon} + \delta\right) < \frac{1-\gamma_0}{b_0-d_0} - \delta \leq \frac{T_{\varepsilon K}^{(0,+)}}{\log(\varepsilon K)} \mid T_0^{(1,-)} = \infty, T_0^{(0,+)} = \infty\right) \\
& \geq \mathbb{P}\left(\frac{T_{\varepsilon K}^{(1,-)}}{\log(\varepsilon K)} \leq \alpha + \frac{1-\gamma_0}{b_1-d_1-(C_{0,1}+C_{1,1})\varepsilon} + \delta \mid T_0^{(1,-)} = \infty\right) \\
& \times \mathbb{P}\left(\frac{T_{\varepsilon K}^{(0,+)}}{\log(\varepsilon K)} \geq \frac{1-\gamma_0}{b_0-d_0} - \delta \mid T_0^{(0,+)} = \infty\right) \\
& \geq \mathbb{P}\left(\frac{T_{\varepsilon K}^{(1,-)}}{\log(\varepsilon K)} \leq \alpha\left(\frac{\log K}{\log(\varepsilon K)}\right) + \frac{1-\gamma_1}{b_1-d_1-(C_{0,1}+C_{1,1})\varepsilon} + \frac{\delta}{2} \mid T_0^{(1,-)} = \infty\right) \\
& \times \mathbb{P}\left(\frac{T_{\varepsilon K}^{(0,+)}}{\log(\varepsilon K)} \geq \frac{1-\gamma_0}{b_0-d_0} - \delta \mid T_0^{(0,+)} = \infty\right)
\end{aligned}$$

Note that in both cases ( $\gamma_1 \geq 0$ ) we have by Lemma 5.1.1,

$$\begin{aligned}
& \mathbb{P}\left(\frac{T_{\varepsilon K}^{(1,-)}}{\log(\varepsilon K)} \leq \alpha\left(\frac{\log K}{\log(\varepsilon K)}\right) + \frac{1-\gamma_1}{b_1-d_1-(C_{0,1}+C_{1,1})\varepsilon} + \frac{\delta}{2} \mid T_0^{(1,-)} = \infty\right) \\
& = \mathbb{P}\left(\frac{T_{\varepsilon K}^{(1,-)}}{\log(\varepsilon K)} \leq \frac{T_D}{\log(\varepsilon K)} + \frac{1-\gamma_1}{b_1-d_1-(C_{0,1}+C_{1,1})\varepsilon} + \frac{\delta}{2} \mid T_0^{(1,-)} = \infty\right) \rightarrow 1
\end{aligned} \tag{5.13}$$

as  $K \rightarrow \infty$ , and by part 3 of Lemma A.1 in [5],

$$\mathbb{P}\left(\frac{T_{\varepsilon K}^{(0,+)}}{\log(\varepsilon K)} \geq \frac{1-\gamma_0}{b_0-d_0} - \delta \mid T_0^{(0,+)} = \infty\right) \rightarrow 1 \tag{5.14}$$

as  $K \rightarrow \infty$ . Combining (5.13) and (5.14) we get

$$\mathbb{P}(T_{\varepsilon K}^{(1,-)} < T_{\varepsilon K}^{(0,+)} \mid T_0^{(0,+)} = \infty, T_0^{(1,-)} = \infty) \rightarrow 1 \tag{5.15}$$

as  $K \rightarrow \infty$ . Note that

$$\mathbb{P}(T_0^{(1,-)} = \infty) = \left(1 - \left(\frac{d_1 + \varepsilon(C_{0,1} + C_{1,1})}{b_1}\right)^{K^{\gamma_1}}\right), \tag{5.16}$$

which clearly converges to 1 if  $\gamma_1 > 0$  and converges to  $1 - (d_1/b_1) - \varepsilon(C_{0,1} + C_{1,1})/d_1$  if  $\gamma_1 = 0$ . Additionally note that as  $\gamma_0 > 0$ ,

$$\mathbb{P}(T_0^{(0,+)} = \infty) = 1 - \left(\frac{d_0}{b_0}\right)^{K^{\gamma_0}} \rightarrow 1 \tag{5.17}$$



as  $K \rightarrow \infty$ . Combining each of the limits (5.16) and (5.17) with the limit (5.15), it follows from (5.12) that

$$\begin{aligned} \mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}) &\rightarrow 1 \quad \text{if } \gamma_1 > 0 \\ \liminf_{K \rightarrow \infty} \mathbb{P}(T_{\varepsilon K}^{(1)} < T_{\varepsilon K}^{(0)}) &\geq 1 - \left( \frac{d_1 + \varepsilon(C_{0,1} + C_{1,1})}{b_1} \right) \quad \text{if } \gamma_1 = 0. \end{aligned} \quad (5.18)$$

(iv)  $\gamma_0 > 0$ ,  $\mathbb{P}(T_{\varepsilon K}^{(0)} < T_{\varepsilon K}^{(1)})$

In view of part (iii), if  $\gamma_1 > 0$  then we immediately have  $\mathbb{P}(T_{\varepsilon K}^{(0)} < T_{\varepsilon K}^{(1)}) \rightarrow 0$  as  $K \rightarrow \infty$ . Assume now that  $\gamma_1 = 0$ . As in part (iii), the fact that  $\gamma_0 > 0$  ensures the type 0 population size stays positive with high probability. Using the couplings for the processes of each type as before,

$$\begin{aligned} \mathbb{P}(T_{\varepsilon K}^{(0)} < T_{\varepsilon K}^{(1)}) &\geq \mathbb{P}\left(T_{\varepsilon K}^{(0)} < T_{\varepsilon K}^{(1)}, T_0^{(0,-)} = \infty\right) \\ &\geq \mathbb{P}\left(T_{\varepsilon K}^{(0,-)} < T_{\varepsilon K}^{(1,+)}, T_0^{(0,-)} = \infty\right) \\ &\geq \mathbb{P}\left(T_{\varepsilon K}^{(0,-)} < T_{\varepsilon K}^{(1,+)}, T_0^{(0,-)} = \infty, T_0^{(1,+)} < T_{\varepsilon K}^{(1,+)}\right) \\ &= \mathbb{P}\left(T_{\varepsilon K}^{(0,-)} < \infty, T_0^{(0,-)} = \infty, T_0^{(1,+)} < T_{\varepsilon K}^{(1,+)}\right) \\ &= \mathbb{P}\left(T_{\varepsilon K}^{(0,-)} < \infty \mid T_0^{(0,-)} = \infty\right) \mathbb{P}\left(T_0^{(0,-)} = \infty\right) \mathbb{P}\left(T_0^{(1,+)} < T_{\varepsilon K}^{(1,+)}\right) \\ &= \mathbb{P}\left(T_0^{(0,-)} = \infty\right) \mathbb{P}\left(T_0^{(1,+)} < T_{\varepsilon K}^{(1,+)}\right). \end{aligned} \quad (5.19)$$

Since  $\gamma_0 > 0$  we have

$$\mathbb{P}\left(T_0^{(0,-)} = \infty\right) = 1 - \left( \frac{d_0 + \varepsilon(C_{0,0} + C_{0,1})}{b_0} \right)^{K\gamma_0} \rightarrow 1 \quad (5.20)$$

as  $K \rightarrow \infty$ , and since  $\gamma_1 = 0$ , by part 1 of Lemma A.1 in [5] we have

$$\mathbb{P}\left(T_0^{(1,+)} < T_{\varepsilon K}^{(1,+)}\right) = \left( 1 - \frac{1 - \frac{d_1}{b_1}}{1 - \left(\frac{d_1}{b_1}\right)^{\varepsilon K}} \right) \rightarrow \frac{d_1}{b_1} \quad (5.21)$$

as  $K \rightarrow \infty$ . Combining the limits (5.20) and (5.21), we have from (5.19):

$$\liminf_{K \rightarrow \infty} \mathbb{P}(T_{\varepsilon K}^{(0)} < T_{\varepsilon K}^{(1)}) \geq \frac{d_1}{b_1}. \quad (5.22)$$

In summary,

$$\begin{aligned} \mathbb{P}(T_{\varepsilon K}^{(0)} < T_{\varepsilon K}^{(1)}) &\rightarrow 0 && \text{if } \gamma_1 > 0 \\ \liminf_{K \rightarrow \infty} \mathbb{P}(T_{\varepsilon K}^{(0)} < T_{\varepsilon K}^{(1)}) &\geq \frac{d_1}{b_1} && \text{if } \gamma_1 = 0. \end{aligned} \tag{5.23}$$

$$\text{(v) } \gamma_0 = 0, \gamma_1 = 0, \mathbb{P}(T_0^{(0)} \vee T_0^{(1)} < T_{\varepsilon K}^{(0)} \wedge T_{\varepsilon K}^{(1)})$$

For this case, just note that

$$\begin{aligned} \mathbb{P}(T_0^{(0)} \vee T_0^{(1)} < T_{\varepsilon K}^{(0)} \wedge T_{\varepsilon K}^{(1)}) &\geq \mathbb{P}(T_0^{(0,+)} < T_{\varepsilon K}^{(0,+)}) \mathbb{P}(T_0^{(1,+)} < T_{\varepsilon K}^{(1,+)}) \\ &= \left(1 - \frac{1 - \frac{d_0}{b_0}}{1 - \left(\frac{d_0}{b_0}\right)^{\varepsilon K}}\right) \left(1 - \frac{1 - \frac{d_1}{b_1}}{1 - \left(\frac{d_1}{b_1}\right)^{\varepsilon K}}\right) \rightarrow \frac{d_0}{b_0} \frac{d_1}{b_1} \end{aligned}$$

as  $K \rightarrow \infty$ . □

The second half of the theorem follows in a similar way.

Now we look at the dynamics during the second phase, which follows straightforwardly from existing results. If one type has reached size  $\varepsilon K$  first, the population size of the other type will be of order  $K^\eta$  for some  $\eta < 1$ . Since the invasion fitness of each type is negative with respect to the other type, it is impossible for the non-established type to go to fixation or invade and reach a coexisting equilibrium. There exists  $M > 0$  such that the type of size  $K^\eta$  at the end of the first phase will belong to the set  $J_{2M\varepsilon}^K = [K^{\eta+2M\varepsilon}, K^{\eta-2M\varepsilon}]$  for the duration of the second phase (with probability converging to 1 as  $K \rightarrow \infty$ ). The following fact is proved as part of Lemma 4.1 of [5] (the key to the proof being the use of Lemma A.2 in [5]):

**Theorem 5.1.3.** *Let  $i \in \{0, 1\}$  denote the first type to reach level  $\varepsilon K$  and let  $j$  denote the other type. For any  $\varepsilon > 0$ , let  $t_\varepsilon(z_i, z_j) > 0$  denote the first time the deterministic system*

$$\begin{aligned} \frac{dn_i}{dt} &= (b_i - d_i - C_{i,i}n_i - C_{i,j}n_j)n_i, && n_i(0) = z_i \\ \frac{dn_j}{dt} &= (b_j - d_j - C_{i,j}n_i - C_{j,j}n_j)n_j, && n_j(0) = z_j \end{aligned}$$

*reaches an  $\varepsilon^2$ -neighbourhood of its stable equilibrium, specifically if we denote this equi-*

librium density by  $\bar{n}_{i,j}$ , define

$$t_\varepsilon(z_i, z_j) = \inf \left\{ t \geq 0 : \sup_{s \in [t, \infty)} \|n_i(s) - \bar{n}_{i,j}\|_1 \leq \varepsilon^2 \right\}.$$

Then

$$\lim_{K \rightarrow \infty} \mathbb{P} \left( \bigcup_{s \in [T_{\varepsilon K}^{(i)}, T_{\varepsilon K}^{(i)} + t_\varepsilon(\varepsilon K, 0)]} \left\{ N_j(s) \in J_{2M\varepsilon}^K \right\}, \left| \frac{N_i(t_\varepsilon(\varepsilon K, 0))}{K} - \bar{n}_i(t_\varepsilon(\varepsilon K, 0)) \right| \leq \varepsilon^2 \right) = 1.$$

*Proof.* Already proved in [5]. □

During the third phase, the resident type remains close to its equilibrium size for a time exponentially large in  $K$ , while the non-resident type eventually goes extinct on a  $\log K$  timescale. This follows since the death rate of the non-resident type (assume it is the type 0 individuals) for some  $\nu > 0$  is at least  $d_0 + \frac{C_{0,0}}{K} K^{1-\nu} + \frac{C_{0,1}}{K} \bar{n}_1 K \approx d_0 + C_{0,1} \bar{n}_1$  for large  $K$ . Due to the negative invasion fitness,  $d_0 + C_{0,1} \bar{n}_1 > b_0$  and hence the population size of the non-resident type evolves approximately as a subcritical birth-death process.

## 5.2 Possible Extension to Allometric Scaling

One approach to extending Theorem 1.2.2 is to consider birth and death rates  $b_i^K$  and  $d_i^K$  for  $i \in \{0, 1\}$  that depend on  $K$ , in such a way that for large  $K$ , births and deaths occur increasingly quickly. In the case of a standard birth-death process, the birth and death rates become  $b_K = b + \eta K^\nu$  and  $d_K = d + \eta K^\nu$  respectively, for some  $\eta, \nu > 0$ . This scaling has been studied in an adaptive dynamics context by Champagnat, Ferrière and Méléard in [9], where it is referred to as the *allometric scaling*, and Theorem A.4 in [9] implies that the population density of a single birth-death process of this type with  $\nu = 1$  (including the case of competition with a logistic term) starting with  $\varepsilon K$  individuals, converges<sup>1</sup> as  $K \rightarrow \infty$  to the solution of a stochastic differential equation, as opposed to the solution of an ordinary differential equation as in the case of birth and death rates independent of  $K$  (for  $\nu < 1$  it is shown that the deterministic limit holds). It should be noted that Theorem A.4 in [9] is a general measure-valued result and thus covers multi-type birth-death processes with density-dependent interactions. In the simplest case of a single linear birth-death process, assume  $b > d$  so the process is supercritical. Then the expected population size at time  $t_K = a \log K$ ,  $a > 0$ , starting

---

<sup>1</sup>For fixed  $T > 0$ , the laws of the trajectories of the processes  $(N_K(t)/K)_{t \geq 0}$  converge weakly with respect to the Skorokhod topology on the space of càdlàg paths  $D_{[0, \infty)}[0, T]$ .

with  $K^\beta$  individuals is

$$\mathbb{E}(N_K(t_K)) = e^{(b_K - d_K)t_K} N_K(0) = e^{(b-d)a \log K} K^\beta = K^{a(b-d)+\beta}$$

which is the same as for the non-allometric scaling. However, the probability of extinction by time  $t_K$ ,  $\mathbb{P}(T_0 \leq t_K)$  is

$$\begin{aligned} \left( \frac{d_K(1 - e^{-(b_K - d_K)t_K})}{b_K - d_K e^{-(b_K - d_K)t_K}} \right)^{N_K(0)} &= \left( \frac{(d + \eta K)(1 - e^{-(b-d)a \log K})}{(b + \eta K) - (d + \eta K)e^{-(b-d)a \log K}} \right)^{K^\beta} \\ &= \left( \frac{(d + \eta K)(1 - K^{-a(b-d)})}{(b + \eta K) - (d + \eta K)K^{-a(b-d)}} \right)^{K^\beta} \\ &= \left( \frac{d - dK^{-a(b-d)} - \eta K^{1-a(b-d)} + \eta K}{b - dK^{-a(b-d)} - \eta K^{1-a(b-d)} + \eta K} \right)^{K^\beta} \\ &\rightarrow 1 \end{aligned}$$

as  $K \rightarrow \infty$  for any  $\beta \in [0, 1)$ . Thus we have an interesting situation in which  $\mathbb{E}(N_K(t_K)) \rightarrow \infty$  as  $K \rightarrow \infty$  however  $\mathbb{P}(N_K(t_K) = 0) = \mathbb{P}(T_0 \leq t_K) \rightarrow 1$  as  $K \rightarrow \infty$ . In our case with two mutant types and density-dependent competition, we would consider a pair of coupled logistic birth-death processes  $(N_0(t), N_1(t))_{t \geq 0}$  such that at time  $t \in [0, \infty)$ ,

- Each type 0 individual gives birth at rate  $b_0 + \eta K$  and dies at rate  $d_0 + \frac{C_{0,0}N_0(t)}{K} + \frac{C_{0,1}N_1(t)}{K} + \eta K$ .
- Each type 1 individual gives birth at rate  $b_1 + \eta K$  and dies at rate  $d_1 + \frac{C_{1,1}N_1(t)}{K} + \frac{C_{0,1}(t)N_0(t)}{K} + \eta K$ .

Note that due to the faster scaling of birth and death rates, trajectories of the approximating linear birth-death processes for  $N_0(t)$  and  $N_1(t)$  will no longer be concentrated with high probability around their respective mean trajectories. To determine whether the type 1 process reaches size  $\varepsilon K$  before the type 0 process, one will need to condition on the size of the type 0 process at time  $T_D = \alpha \log K$  and use the full distribution of  $N_0(T_D)$ . Additionally, the existing results for the asymptotic time taken to reach level  $K^v$  from an initial size of  $K^u$  for  $0 \leq u < v \leq 1$  will no longer be valid in this context (as the accelerated birth and death rates will no longer allow an almost-sure limit for large  $K$ ) so the full distributions of the times  $T_{\varepsilon K}^{(0)}$  and  $T_{\varepsilon K}^{(1)}$  will have to be taken into account.

# Appendix A

## Appendix

In Section A.1 we outline the construction of the stochastic integral with respect to Poisson point processes and quantities related to such integrals. The material follows closely the survey paper of Bass [3]. In Section A.2 we recall Itô's formula for general semimartingales and in Section A.3 prove a result needed for Example 1.1.9 in the Introduction.

### A.1 Stochastic Integral With Respect To Poisson Processes

Recall that if  $M$  is a càdlàg square-integrable martingale, then  $(\langle M \rangle_t)_{t \geq 0}$  is the unique increasing predictable process such that  $M_t^2 - \langle M \rangle_t$  is a martingale. Additionally define

$$[M]_t = \langle M^c \rangle_t + \sum_{s \leq t} |\Delta M_s|^2,$$

where  $M^c$  is the continuous part of  $M$  (see Theorem 2.1 in [3]) and  $\Delta M_s = M_s - M_{s-}$  is the process recording the jumps of  $M$ . The process  $[M]_t$  is known as the *quadratic variation* while  $\langle M \rangle_t$  is known as the *predictable quadratic variation*. Let  $(S, \Sigma, \xi)$  be some measure space and let  $\mu$  be a Poisson random measure on  $[0, \infty) \times S$  of intensity measure  $\nu = \lambda \otimes \xi$  where  $\lambda$  denotes Lebesgue measure. If  $\xi(A) < \infty$ , the process  $(\eta_t)_{t \geq 0}$  defined by  $\eta_t = \mu([0, t] \times A) - \nu([0, t] \times A)$  is the difference between a Poisson process and its mean, and consequently is a martingale.

The following construction of the integral with respect to a Poisson process can be found in Section 3 of [3]. By starting with elementary predictable functions  $H$ , and

then extending to general predictable functions such that  $\int_{[0,t] \times S} H^2(s, z) \nu(s, z) < \infty$  almost surely for every  $t > 0$ , one can define the general stochastic integral

$$N_t = \int_0^t \int_S H(s, z) (\mu - \nu)(ds, dz).$$

Moreover, this process has the property that

$$[N]_t = \int_0^t \int_S H(s, z)^2 \mu(ds, dz)$$

and

$$\langle N \rangle_t = \int_0^t \int_S H(s, z)^2 \nu(ds, dz).$$

We now introduce an example relevant for the processes considered in Chapter 4.

**Example:** Let  $b = (b_t)_{t \geq 0}$  and  $d = (d_t)_{t \geq 0}$  be measurable, deterministic functions. We can write the  $(b, d)$ -branching process as

$$N_t = N_0 + \int_{[0,t] \times [0,\infty)} \left( \mathbb{1}_{[0, b_s N_{s-}]}(z) - \mathbb{1}_{[b_s N_{s-}, (b_s + d_s) N_{s-}]}(z) \right) d\mu(s, z),$$

where  $\mu$  is a unit-intensity Poisson random measure on  $[0, \infty)^2$ . In this case, the intensity measure  $\nu$  is just Lebesgue measure on  $[0, \infty)$ . In particular, we can write  $N_t$  for each  $t \geq 0$  in the compensated form as

$$\begin{aligned} N_t &= N_0 + M_t + \int_0^t \int_0^\infty \left( \mathbb{1}_{[0, b_s N_{s-}]}(z) - \mathbb{1}_{[b_s N_{s-}, (b_s + d_s) N_{s-}]}(z) \right) ds dz \\ &= N_0 + M_t + \int_0^t (b_s N_{s-} - d_s N_{s-}) ds, \end{aligned}$$

where  $M_t$  is defined via

$$M_t := \int_{[0,t] \times [0,\infty)} \left( \mathbb{1}_{[0, b_s N_{s-}]}(z) - \mathbb{1}_{[b_s N_{s-}, (b_s + d_s) N_{s-}]}(z) \right) d(\mu - \nu)(s, z).$$

Then, from the general theory we can read off that the predictable quadratic variation

is given as

$$\begin{aligned}
\langle M \rangle_t &= \int_{[0,t] \times [0,\infty)} \left( \mathbb{1}_{[0,b_s N_{s-}]}(z) - \mathbb{1}_{[b_s N_{s-}, (b_s+d_s) N_{s-}]}(z) \right)^2 ds dz \\
&= \int_{[0,t] \times [0,\infty)} \left( \mathbb{1}_{[0,b_s N_{s-}]}(z) + \mathbb{1}_{[b_s N_{s-}, (b_s+d_s) N_{s-}]}(z) \right) ds dz \\
&= \int_0^t (b_s + d_s) N_s ds.
\end{aligned}$$

## A.2 Itô's Formula for Jump Processes

Let  $(X_t)_{t \geq 0}$  be a real-valued semimartingale and let  $(X_t^c)_{t \geq 0}$  denote its continuous part. For  $t > 0$ , we use the notation  $X_{t-} = \lim_{s \nearrow t} X_s$ . Then for all  $f \in C^{1,2}([0, \infty) \times \mathbb{R})$  and  $t \geq 0$ , Itô's formula states

$$\begin{aligned}
f(t, X_t) &= f(0, X_0) + \int_0^t \frac{\partial f}{\partial s}(s, X_s) ds + \int_0^t \frac{\partial f}{\partial x}(s, X_s) dX_s + \frac{1}{2} \int_0^t \frac{\partial^2 f}{\partial x^2}(s, X_s) d\langle X^c \rangle_s \\
&\quad + \sum_{0 \leq s \leq t} \left( f(s, X_s) - f(s, X_{s-}) - \frac{\partial f}{\partial s}(s, X_{s-})(X_s - X_{s-}) \right).
\end{aligned}$$

Next we introduce an example relevant for the processes considered in Chapter 4.

**Example:** Let  $(X_t)_{t \geq 0}$  be a real-valued semimartingale and suppose  $(M_t)_{t \geq 0}$  is a martingale such that for each  $t \geq 0$ ,

$$M_t = X_t - X_0 + \int_0^t g(s) X_s ds$$

for some integrable function  $g(t)$ . Set  $h(t) = \int_0^t g(s) ds$  and define  $Y_t = e^{h(t)} X_t$ . For  $f(t, x) = e^{h(t)} x$ , we have  $\frac{\partial f}{\partial s}(s, X_s) = h'(s) e^{h(s)} X_s$ ,  $\frac{\partial f}{\partial x}(s, X_s) = e^{h(s)}$  and  $\frac{\partial^2 f}{\partial x^2}(s, X_s) = 0$ . Hence,

$$\begin{aligned}
Y_t = f(t, X_t) &= e^{h(0)} X_0 + \int_0^t h'(s) e^{h(s)} X_s ds + \int_0^t e^{h(s)} dX_s \\
&\quad + \sum_{0 \leq s \leq t} \left( e^{h(s)} X_s - e^{h(s)} X_{s-} - e^{h(s)} (X_s - X_{s-}) \right) \\
&= X_0 + \int_0^t g(s) e^{h(s)} X_s ds + \int_0^t e^{h(s)} dX_s \\
&= X_0 + \int_0^t e^{h(s)} dM_s.
\end{aligned}$$

### A.3 Factorial Branching Number on Spherically Symmetric Trees

Recall that for a tree  $\mathcal{T}$  with set of cutsets  $\Pi$ , the factorial branching number  $\text{br}_f(\mathcal{T})$  is defined by

$$\text{br}_f(\mathcal{T}) = \sup \left\{ \lambda > 0 : \inf_{\pi \in \Pi} \sum_{e \in \pi} \left( \frac{1}{|e|!} \right)^\lambda > 0 \right\}.$$

In this section the aim is to show that for spherically symmetric trees, the infimum over cutsets in the right hand side of the above expression can be replaced by the infimum over sets of edges at each level. We will use Theorem 3.1 in [29] and to state it we need some concepts first introduced in Section 3.2. For each edge  $e \in E$  let  $D_e$  denote the set of descendant edges at the next level. A function  $\theta : E \rightarrow [0, \infty)$  is called a *flow* if for all  $e \in E$ ,

$$\theta(e) = \sum_{d \in D_e} \theta(d).$$

For any flow  $\theta$  on  $E$ , the *strength* of  $\theta$  is given by

$$\text{Strength}(\theta) = \sum_{e \in E: |e|=1} \theta(e) \tag{A.1}$$

and note that  $\text{Strength}(\theta)$  can equal zero. We will also refer to a given fixed function  $c : E \rightarrow [0, \infty)$ , viewed as a collection of values  $(c(e))_{e \in E}$ , as an assignment of *conductances* to the edges in  $E$ .

The *Max-Flow Min-Cut Theorem*, which in the version for infinite networks is stated as Theorem 3.1 in [citeLP], can be re-stated in our context as

**Theorem A.3.1** (Theorem 3.1 in citeLP). *For a locally finite tree  $\mathcal{T} = (V, E)$  and conductances  $(c(e))_{e \in E}$  on the edges of  $\mathcal{T}$ , let  $\Theta \in [0, \infty)^E$  denote the set of flows  $\theta$  on  $E$  such that  $0 \leq \theta(e) \leq c(e)$  for all  $e \in E$ . Let  $\Pi$  denote the set of cutsets of  $\mathcal{T}$ . Then*

$$\max_{\theta \in \Theta} \left\{ \text{Strength}(\theta) \right\} = \inf_{\pi \in \Pi} \sum_{e \in \pi} c(e).$$

We use this to prove the following proposition:

**Proposition A.3.2.** *Let  $\mathcal{T} = (V, E)$  be a spherically symmetric tree such that each*



vertex at level  $h \geq 0$  has degree  $d(h)$ . Then

$$\text{br}_f(\mathcal{T}) = \sup \left\{ \lambda > 0 : \inf_{h \geq 1} \sum_{|e|=h} \left( \frac{1}{|e|!} \right)^\lambda > 0 \right\}. \quad (\text{A.2})$$

*Proof.* First note that by definition

$$\text{br}_f(\mathcal{T}) \leq \sup \left\{ \lambda > 0 : \inf_{h \geq 1} \sum_{|e|=h} \left( \frac{1}{|e|!} \right)^\lambda > 0 \right\}, \quad (\text{A.3})$$

since the set of edges separating level  $h$  from level  $h+1$  is a cutset for all  $h \geq 0$ . The aim is now to show that equality holds in (A.3). For each  $\lambda \geq 0$ , we can define conductances via  $c_\lambda(e) = (|e|!)^{-\lambda}$ . Let  $\Theta(\lambda)$  be the set of flows such that  $0 \leq \theta(e) \leq c_\lambda(e)$  as required for Theorem A.3.1. The factorial branching number can then be expressed as

$$\begin{aligned} \text{br}_f(\mathcal{T}) &= \sup \left\{ \lambda > 0 : \inf_{\pi \in \Pi} \sum_{e \in \pi} \left( \frac{1}{|e|!} \right)^\lambda > 0 \right\} \\ &= \sup \left\{ \lambda > 0 : \max_{\theta \in \Theta(\lambda)} \left\{ \text{Strength}(\theta) \right\} > 0 \right\} \end{aligned} \quad (\text{A.4})$$

Now suppose

$$\hat{\lambda} = \sup \left\{ \lambda > 0 : \inf_{h \geq 1} \sum_{|e|=h} \left( \frac{1}{|e|!} \right)^\lambda > 0 \right\} > 0.$$

Then there exists a value  $\alpha = \alpha(\hat{\lambda}) > 0$  such that

$$\inf_{h \geq 1} \sum_{|e|=h} \left( \frac{1}{|e|!} \right)^{\hat{\lambda}} = \alpha.$$

The aim is now to exhibit a nonzero flow in  $\Theta(\hat{\lambda})$ . Choose  $\beta \in (0, \alpha]$  and consider the flow  $\hat{\theta}$  on  $E$  such that when  $|e| = n$  for any  $n \geq 1$ ,

$$\hat{\theta}(e) = \frac{\beta}{\prod_{i=1}^n d(i)}.$$

Observe that by spherical symmetry the set  $\{e \in E : |e| = n\}$  has size  $\prod_{i=1}^n d(i)$  and

hence

$$\hat{\theta}(e) \leq \frac{\alpha}{\prod_{i=1}^n d(i)} = \frac{\inf_{h \geq 1} \sum_{|e|=h} \left(\frac{1}{|e|!}\right)^{\hat{\lambda}}}{\prod_{i=1}^n d(i)} \leq \frac{\sum_{|e|=n} \left(\frac{1}{|e|!}\right)^{\hat{\lambda}}}{\prod_{i=1}^n d(i)} = \left(\frac{1}{|e|!}\right)^{\hat{\lambda}}.$$

It follows that  $\hat{\theta}$  is a nonzero flow on  $E$  (in particular of strength  $\beta > 0$ ) satisfying  $0 \leq \hat{\theta}(e) \leq c_{\hat{\lambda}}(e)$ . Therefore  $\hat{\theta} \in \Theta(\hat{\lambda})$  which implies

$$\max_{\theta \in \Theta(\hat{\lambda})} \left\{ \text{Strength}(\theta) \right\} \geq \beta > 0.$$

In view of (A.4) we then have  $\hat{\lambda} \leq \text{br}_f(\mathcal{T})$ . Combining with (A.3) implies that equation (A.2) holds.  $\square$

# Bibliography

- [1] Atish Agarwala and Daniel S. Fisher. Adaptive walks on high-dimensional fitness landscapes and seascapes with distance-dependent statistics. *Theoretical Population Biology*, 130:13–49, 2019.
- [2] K.B. Athreya and P.E. Ney. *Branching processes*. Dover Publications, 2004.
- [3] Richard F. Bass. Stochastic differential equations with jumps. *Probability Surveys*, 1:1–19, 2004.
- [4] Necdet Batir. Sharp inequalities for factorial. *Proyecciones (antofagasta)*, 27, 05 2008.
- [5] Sylvain Billiard and Charline Smadi. The interplay of two mutations in a population of varying size: A stochastic eco-evolutionary model for clonal interference. *Stochastic Processes and their Applications*, 127(3):701–748, 2017.
- [6] Niloy Biswas, A. Etheridge, and Aleksander Klimek. The spatial Lambda-Fleming-Viot process with fluctuating selection. *Electronic Journal of Probability*, 26, 02 2018.
- [7] David Branson. Inhomogeneous birth-death and birth-death-immigration processes and the logarithmic series distribution. *Stochastic Processes and their Applications*, 39:131–137, 10 1991.
- [8] Nicolas Champagnat. A microscopic interpretation for adaptive dynamics trait substitution sequence models. *Stochastic Processes and their Applications*, 116(8):1127–1160, 2006.
- [9] Nicolas Champagnat, Régis Ferrière, and Sylvie Méléard. Unifying evolutionary dynamics: from individual stochastic processes to macroscopic models. *Theoretical population biology*, 69 3:297–321, 2006.
- [10] Nicolas Champagnat, Sylvie Meleard, and Viet Chi Tran. Stochastic analysis of

- emergence of evolutionary cyclic behavior in population dynamics with transfer. *The Annals of Applied Probability*, 31, 08 2021.
- [11] Xinxin Chen. Increasing paths on  $n$ -ary trees. *Preprint, arXiv:1403.0843*, 2014.
- [12] Cristian F. Coletti, Renato J. Gava, and Pablo M. Rodriguez. On the existence of accessibility in a tree-indexed percolation model. *Physica A: Statistical Mechanics and its Applications*, 492:382–388, 2018.
- [13] Andrea Collecchio, Cong Bang Huynh, and Daniel Kious. The branching-ruin number as critical parameter of random processes on trees. *Electronic Journal of Probability*, 24:1 – 29, 2019.
- [14] L. Chao D. M. Weinreich, R. A. Watson. Perspective: Sign epistasis and genetic constraint on evolutionary trajectories. *Evolution*, 59(6):1165–74, jun 2005.
- [15] Suman Das, Joachim Krug, and Muhittin Mungan. A driven disordered systems approach to biological evolution in changing environments. *Preprint, arXiv:2108.06170*, 2021.
- [16] Frank Duque, Alejandro Roldn-Correa, and Leon Valencia. Accessible percolation with crossing valleys on  $n$ -ary trees. *Journal of Statistical Physics*, 174, 03 2019.
- [17] S.N. Ethier and T.G. Kurtz. *Markov Processes: Characterization and Convergence*. John Wiley Sons, Inc., 2005.
- [18] R. A. Fisher. *The Genetical Theory of Natural Selection*. Clarendon Press, 1930.
- [19] Nicolas Fournier and Sylvie Meleard. A microscopic probabilistic description of a locally regulated population and macroscopic approximations. *Annals of Applied Probability*, 14, 11 2004.
- [20] C. Garban, G. Pete, and O. Schramm. The Fourier spectrum of critical percolation. *Acta Mathematica*, 205(1):19–104, 2010.
- [21] James S. Horton, Louise M. Flanagan, Robert W. Jackson, Nicholas K. Priest, and Tiffany B. Taylor. Extreme parallel evolution of flagellar motility facilitated by silent mutations. *bioRxiv*, 2021.
- [22] J. E. Huxley. *Evolution: The Modern Synthesis*. MIT Press, 2009.
- [23] Svante Janson, Tomasz Luczak, and Andrzej Rucinski. *Random Graphs*. John Wiley Sons, 2000.

- [24] David G. Kendall. On the generalized "birth-and-death" process. *The Annals of Mathematical Statistics*, 19(1):1–15, 1948.
- [25] Davar Khoshnevisan. Dynamical percolation on general trees. *Probab. Theory Relat. Fields*, 140:169–193, 2008.
- [26] Thomas G. Kurtz. Averaging for martingale problems and stochastic approximation. In Ioannis Karatzas and Daniel Ocone, editors, *Applied Stochastic Analysis*, pages 186–209, Berlin, Heidelberg, 1992. Springer Berlin Heidelberg.
- [27] Benedikt Kpfer and Ludger Rschendorf. The martingale comparison method for Markov processes. *Journal of Applied Probability*, 58(1):164–176, 2021.
- [28] Amaury Lambert. The branching process with logistic growth. *The Annals of Applied Probability*, 15(2):1506–1535, 2005.
- [29] Russell Lyons and Yuval Peres. *Probability on Trees and Networks*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2017.
- [30] S. Nowak and J. Krug. Accessibility percolation on n-trees. *Europhysics Letters*, 101(6), apr 2013.
- [31] Ryan O’Donnell. *Analysis of Boolean Functions*. Cambridge University Press, 2014.
- [32] Hggstrm Olle, Peres Yuval, and E. Steif Jeffrey. Dynamical percolation. *Annales de l’Institut Henri Poincare (B) Probability and Statistics*, 33(4):497–528, 1997.
- [33] Matthew Roberts and Bati Şengül. Exceptional times of the critical dynamical Erdős-Rényi graph. *Ann. Appl. Probab.*, 28(4):2275–2308, 2018.
- [34] Matthew Roberts and Lee Zhuo Zhao. Increasing paths in regular trees. *Electronic Communications in Probability*, 18, 2013.
- [35] Oded Schramm and J. E. Steif. Quantitative noise sensitivity and exceptional times for percolation. *Annals of Mathematics*, 171(2):619–672, 2010.
- [36] J E Steif. A survey of dynamical percolation. In C Bandt, M Zähle, and P Mörters, editors, *Fractal Geometry and Stochastics IV*, volume 61 of *Progress in Probability*. Birkhäuser Basel, 2009.
- [37] Claus O. Wilke and Thomas Martinetz. Adaptive walks on time-dependent fitness landscapes. *Phys. Rev. E*, 60:2154–2159, Aug 1999.

- [38] S. G. Wright. Evolution in Mendelian populations. *Genetics*, 16(2):97–159, mar 1931.