Electronic Journal of Probability

Electron. J. Probab. **26** (2021), article no. 151, 1–44.
ISSN: 1083-6489 https://doi.org/10.1214/21-EJP704

The Wright–Fisher model for class–dependent fitness landscapes

Joseba Dalmau*

Abstract

We consider a population evolving under mutation and selection. The genotype of an individual is a word of length ℓ over a finite alphabet. Mutations occur during reproduction, independently on each locus; the fitness depends on the Hamming class (the distance to a reference sequence w^*). Evolution is driven according to the classical Wright–Fisher process. We focus on the proportion of the different classes under the invariant measure of the process. We consider the regime where the length of the genotypes ℓ goes to infinity, and

$$\text{population size } \sim \ell, \qquad \text{mutation rate } \sim 1/\ell.$$

We prove the existence of a critical curve, which depends both on the population size and the mutation rate. Below the critical curve, the proportion of any fixed class converges to 0, whereas above the curve, it converges to a positive quantity, for which we give an explicit formula.

Keywords: Wright–Fisher model; quasispecies; invariant measure; error threshold; large deviations.
MSC2020 subject classifications: 60J10.
Submitted to EJP on April 20, 2020, final version accepted on September 7, 2021.
Supersedes arXiv:1712.00279v1.

1 Introduction

Most of the living populations share, among others, these three main features: genomes are long, populations are large, and mutations are rare. Nevertheless, when modeling a living population, different relations between those three parameters will lead to different conclusions. We focus here on a situation which is most appropriate for living beings of small complexity, as RNA viruses, or replicating macromolecules: we aim to model a population in which both the population size and the inverse mutation rate are of the same order as the length of the genome [10]. The main forces that will drive the evolution of such a population are, of course, mutation, but also selection, and genetic

*NYU–ECNU Institute of Mathematical Sciences at NYU Shanghai. E-mail: dalmau.joseba@gmail.com

drift. Selection is introduced via a fitness function on the genotypes, which encodes the average number of offspring of an individual carrying a particular genotype. Genetic drift is introduced by considering a finite population of constant size. This modeling situation is known to lead to very particular and interesting phenomena:

Error threshold. There is a critical mutation rate separating two different regimes. Above the critical mutation rate, all genetic information is eventually lost, while below the critical mutation rate, an equilibrium state is reached in which the fittest genotype (the master sequence) is present in a positive proportion.

Quasispecies. The equilibrium that is reached below the error threshold consists of a positive proportion of the fittest genotype, which may be very low, and mutants that are a few mutations away from the master sequence may appear in high proportions. Thus, the genetic heterogeneity of such an equilibrium state is huge, and we might as well not be able to identify the master sequence. Such a population is often referred to as a quasispecies.

Population threshold. A low mutation rate is not enough for a quasispecies to form. Indeed, if the population is too small, it is likely that the master sequences present in the population mutate all at once or in a few generations, thus loosing the driving force of the quasispecies. This event becomes more and more unlikely as the population size grows, thus giving rise to a second threshold phenomenon, namely a population threshold.

The first two phenomena where first observed by Eigen, in a mathematical model for prebiotic populations [7]. The concept of quasispecies was later popularized by Eigen and Schuster [8]. The model considered by Eigen takes the population size to be infinite, and models the evolution via a system of differential equations. The system is studied in the long chain regime, i.e., when the length of the genomes goes to infinity. It is in this regime that the error threshold and quasispecies phenomena arise. In order to observe the population threshold, it is necessary to consider a model where the population is finite. This phenomenon has first been observed in [1] for the Moran model and [2] for the Wright–Fisher model. A nice account of the error threshold and quasispecies phenomena, the main models where they arise, and their applications can be found in [6]. We refer the reader to [1] for a more detailed exposition of the different attempts to build finite population models that present the error threshold and the quasispecies.

Most of the works that show the above three phenomena deal with the simplest possible fitness landscape, namely the sharp–peak landscape: there is a single fittest genotype, the master sequence, and all the other genotypes share the same fitness. The works [2, 4] show how, in the sharp–peak landscape, the Wright–Fisher model presents all three of the above phenomena. Our objective is to extend these results to more general fitness landscapes. We focus in the present paper on the case of class–dependent fitness functions: there is a single fittest genotype, and the fitness of any other genotype is a function of its Hamming distance to the fittest genotype. We present the model in Section 2.1, while the main result is presented in Section 2.6, along with a sketch of the proof in Section 2.8. The remaining sections are devoted to the proof of the main result.

2 Model and results

2.1 Genotypes, fitness and mutation

Let $\mathcal{A}$ be a finite alphabet of cardinality $\kappa \geq 2$, and let $\ell \geq 1$ represent the length of the genome. We consider individuals whose genotypes are elements of $\mathcal{A}^\ell$. Each genotype $u \in \mathcal{A}^\ell$ has a fitness $\hat{A}(u)$ associated to it, which should be interpreted as the mean number of children of an individual carrying the genotype u. When a reproduction occurs, the newborn child is subject to mutations. We suppose that mutations happen independently over each site of the genotype, with probability $q \in (0,1)$. When a particular site mutates, the present letter is replaced with a uniformly chosen letter from the $\kappa - 1$ remaining ones. Thus, the probability of mutating from a chain u to another chain v is given by

$$\hat{M}(u,v) := \left(\frac{q}{\kappa-1}\right)^{d(u,v)} (1-q)^{\ell-d(u,v)},$$

where $d(u,v)$ represents the Hamming distance between u and v, or equivalently, the number of digits the two sequences differ in.

The evolution will be guided by the classical Wright–Fisher process. Nevertheless, the analysis of the Wright–Fisher process for an arbitrary fitness function $\hat{A}$ is far too complicated. We focus here on fitness functions of a particular form, namely the class–dependent fitness functions. We make the following assumptions on $\hat{A}$.

Master sequence. We assume the existence of a genotype $w^* \in \mathcal{A}^\ell$ with maximal fitness, which we call the master sequence.

Class–dependence. We assume further that the fitness of a genotype u depends only on the number of point mutations away from the master sequence. In other words, all the sequences at Hamming distance k from the master sequence form the Hamming class k, and they all share the same fitness.

Eventually constant. Finally we assume that there is a Hamming class $K \geq 0$ such that the genotypes in the classes beyond K have fitness 1.

The idea behind this assumptions is that the master sequence is the most adapted genotype to the actual environment (which is constant), and that fitness decreases with accumulating errors, up to a certain number of errors K. Once this number of errors is reached, accumulating errors changes the fitness by so little that we approximate all the fitnesses of all the subsequent Hamming classes by 1. We do not assume that the fitness function is decreasing, because our main result and proofs work in the more general case where the fitness is not necessarily decreasing. The genotypes carrying more than K errors will be called *neutral genotypes*, and the genotypes carrying K or less errors will be called *non–neutral*. Under the above assumptions, we can define:

Definition 2.1. *Let $A : \mathbb{N} \to \mathbb{R}_+$ be the function such that for all $u \in \mathcal{A}^\ell$ we have $\hat{A}(u) =: A\big(d(u,w^*)\big)$.*

Note that the function A satisfies:
- $A(0) > A(k)$ for all $k \geq 1$.
- $A(K) \neq 1$ and $A(k) = 1$ for all $k \geq K+1$.

When $K = 0$, all the genotypes other than the master sequence have fitness 1. This particular case is referred to as the *sharp–peak landscape*; the Wright–Fisher model on the sharp peak landscape has been studied in detail in [2, 4]. Our aim is to generalize the results therein to class–dependent fitness functions which are eventually constant. One of the main advantages of working with class–dependent fitness functions is that we can break the space $\mathcal{A}^\ell$ into Hamming classes (sets of sequences sharing the same distance to w^*). This is possible because the mutation matrix $\hat{M}$ respects the Hamming classes (cf. the Lemma 6.1 in [1] for a proof): fix $0 \leq k, l \leq \ell$ and let $X \sim Bin(k, q/(\kappa-1))$, $Y \sim Bin(\ell-k, q)$ be independent random variables, then for any $u \in \mathcal{A}^\ell$ in the class k,

$$M(k,l) := \sum_{v:d(v,w^*)=l} \hat{M}(u,v) = P(k - X + Y = l). \tag{2.1}$$

We call M the *lumped mutation matrix*, and A the *lumped fitness function*.

2.2 The Wright–Fisher model

We consider a population of size $m \geq 1$ evolving according to the classical Wright–Fisher process with mutation and selection. Informally, the transition from the population at time n, to the population at time $n+1$ is done as follows: m individuals are sampled from the population at time n, with replacement. At each of the m trials, the probability for a given individual to be chosen is

$$\frac{\text{fitness of the individual}}{\text{sum of all fitnesses in the population}}\,.$$

Each of the m chosen individuals reproduces, giving birth to one child each, and the offspring mutate. The ensemble of the m offspring, after mutation, form the population at time $n+1$. We will only be interested in the proportions of the different Hamming classes, and not on the distribution of the different genotypes inside the classes themselves; the only information we actually need about the population at time n, is the number of individuals in each of the Hamming classes. Indeed, this information is enough to determine the number of individuals in each class at time $n+1$. The process that keeps this information is the occupancy process $(O_n)_{n\geq 0}$ (which we will define shortly) and it will be the starting point of our study. It is obtained from the original Wright–Fisher process by using a technique known as lumping; for a formal definition of the original Wright–Fisher process, as well as for a formal derivation of the occupancy process from it, we refer the reader to Sections 2 and 4 of [2]. Let $\mathcal{P}^m_{\ell+1}$ be the set of the ordered partitions of the integer m in at most $\ell+1$ parts:

$$\mathcal{P}^m_{\ell+1} \,:=\, \big\{\,(o(0),\dots,o(\ell))\in\mathbb{N}^{\ell+1} : o(0)+\cdots+o(\ell)=m\,\big\}\,, \tag{2.2}$$

(we assume that $0\in\mathbb{N}$). A partition $(o(0),\dots,o(\ell))$ is interpreted as an occupancy distribution, i.e. a population with $o(l)$ individuals in the Hamming class $l=0,\dots,\ell$. Note that the number of Hamming classes is equal to the length of the genome, hence the use of the parameter ℓ here.

Definition 2.2. *The occupancy process $(O_n)_{n\geq 0}$ is a Markov chain with values in $\mathcal{P}^m_{\ell+1}$ and transition matrix given by: for $o,o'\in\mathcal{P}^m_{\ell+1}$,*

$$p_O(o,o') \,:=\, \frac{m!}{o'(0)!\cdots o'(\ell)!}\prod_{0\leq h\leq\ell}\left(\frac{\sum_{k\in\{0,\dots,\ell\}} o(k)A(k)M(k,h)}{o(0)A(0)+\cdots+o(\ell)A(\ell)}\right)^{o'(h)}$$

Let $\mathcal{S}^\ell$ denote the ℓ–dimensional unit simplex

$$\mathcal{S}^\ell \,:=\, \big\{\,x\in[0,1]^{\ell+1} : |x|_1=1\,\big\}\,.$$

We define the function $F:\mathcal{S}^\ell\longrightarrow\mathcal{S}^\ell$ by setting

$$\forall\,x\in\mathcal{S}^\ell\quad\forall\,k\in\{\,0,\dots,\ell\,\}\qquad F_k(x)\,:=\,\frac{\displaystyle\sum_{0\leq h\leq\ell} x_hA(h)M(h,k)}{1+\displaystyle\sum_{0\leq h\leq K} x_h(A(h)-1)} \tag{2.3}$$

In view of the expression of the transition matrix, for all $o\in\mathcal{P}^m_{\ell+1}$ and $n\geq 0$, given that $O_n=o$, the random vector O_{n+1} follows a multinomial law with parameters m and $F(o/m)$.

Notation. The expression appearing in the denominator of the function $F(x)$ represents the mean fitness of the population x. Since it will recurrently appear in the subsequent formulas, for any $k\geq K$ and $x\in\mathbb{R}^{k+1}$, we denote

$$\phi(x)\,:=\,1+\sum_{0\leq h\leq K} x_h(A(h)-1)\,. \tag{2.4}$$

2.3 Asymptotic regime

A straightforward treatment of the occupancy process is hardly tractable. Lucky for us, in most living populations, genomes are long, populations large, and mutations rare. We will thus carry out the study of the occupancy process by sending the length of the genomes ℓ and the population size m to infinity, and sending the mutation probability q to 0, as follows:

$$\begin{gathered}\ell \to +\infty\,, \qquad m \to +\infty\,, \qquad q \to 0\,,\\ \ell q \to a \in]0,+\infty[\ , \qquad \frac{m}{\ell} \to \alpha \in]0,+\infty[\ .\end{gathered} \tag{2.5}$$

The parameter a represents the mean number of mutations per genome per reproduction cycle, while α can be thought of as a rescaled population size.

Notation. In the sequel, when the notation $\lim_{m,\ell,q}$ appears, we will mean that we take the limit in the asymptotic regime (2.5). Likewise, we will denote by $\lim_{\ell,q}$ the limit

$$\ell \to +\infty\,, \qquad q \to 0\,, \qquad \ell q \to a\,. \tag{2.6}$$

The asymptotic regime (2.5) has two main consequences on the normalized occupancy process $(O_n/m)_{n\geq 0}$,

• Since $m \to \infty$, the multinomial law involved in the transition mechanism of the process concentrates around its mean, which is given by the mapping F, and the trajectories of the process tend to be close to those of the discrete dynamical system given by the iterations of F.

• Under (2.6), the mutation matrix M converges to an infinite upper diagonal matrix M_∞; the probability of mutating to a lower class converges to 0, and the probability of jumping forward converges to a Poisson law of parameter a (cf. Lemma 2.3 in the next section).

2.4 Properties of the mutation matrix

In order to clarify how the mutation matrix behaves in the asymptotic regime, as well as for further reference, we state here the properties of the mutation matrix $(M(i,j), 0 \leq i,j \leq \ell)$ that are relevant to our case. The i–th row of the lumped mutation matrix is given by the difference of two independent binomial laws, i.e., if $X \sim Bin(i, q/(\kappa-1))$ and $Y \sim Bin(\ell - i, q)$ are independent random variables, then

$$M(i,j) \,=\, P(i - X + Y = j)\,.$$

Fix i and j and let ℓ go to infinity, q go to 0, and ℓq go to a; the first of the binomial laws converges to a Dirac mass at 0, while the second one converges to a Poisson random variable of parameter a. This is summarized in the following lemma.

Lemma 2.3. *For every $i, j \geq 0$, we have the following convergence:*

$$\lim_{\substack{\ell\to\infty,\, q\to 0\\ \ell q\to a}} M(i,j) \,=\, M_\infty(i,j) \,:=\, \begin{cases} e^{-a}\dfrac{a^{j-i}}{(j-i)!} & \textit{if } j \geq i\,,\\ 0 & \textit{otherwise}\,.\end{cases}$$

In particular, in the limit, there is no back mutation.

Let us fix $k \geq 0$. For ℓ large enough, q small enough and ℓq close enough to a, we have

$$M(k+2,k) \,\geq\, M(j,k) \qquad \text{for any } j \geq k+2\,.$$

Moreover, under (2.6),

$$\lim_{\ell,q} M(k+1,k) \,=\, 0 \qquad \text{and} \qquad \lim_{\ell,q} \frac{M(k+2,k)}{\ell} \,=\, 0\,.$$

In particular,

$$\lim_{\ell,q} \sum_{k+1\leq h\leq \ell} M(h,k) \,=\, 0\,. \tag{2.7}$$

2.5 Associated dynamical system

As stated above, when the population is large, the Wright–Fisher process can be viewed as a perturbation of the dynamical system given by the iterations of the mapping F (cf. the formula (2.3)). Nevertheless, in our asymptotic regime (2.5), not only the population is large. The convergence of the mutation matrix M to its infinite version M_∞ (cf. the Lemma 2.3) has an important impact on the dynamical system associated to F. Let $k \geq K$ and define the set $\mathcal{D}^k$ by

$$\mathcal{D}^k \,:=\, \big\{\, r\in [0,1]^{k+1} : |r|_1 \leq 1 \,\big\}\,.$$

The first $k+1$ coordinates of F converge to a mapping $G:\mathcal{D}^k \longrightarrow \mathcal{D}^k$ given by: for $r\in\mathcal{D}^k$ and $i\in\{\,0,\dots,k\,\}$,

$$G_i(r) \,:=\, \phi(r)^{-1}\sum_{h=0}^{i} r_h A(h) e^{-a}\,\frac{a^{i-h}}{(i-h)!}\,. \tag{2.8}$$

Lemma 2.4. *For every $r\in\mathcal{D}^k$ and $x\in\mathcal{S}^\ell$ satisfying $(x_0,\dots,x_k)=(r_0,\dots,r_k)$ we have the convergence*

$$\lim_{\substack{\ell\to\infty,\,q\to 0\\ \ell q\to a}} \big(F_1(x),\dots,F_k(x)\big) \,=\, G(r)\,. \tag{2.9}$$

The above follows from the convergence of the mutation matrix combined with the facts (cf. formula (2.7))

$$\sum_{0\leq h\leq \ell} x_h \,=\, 1 \qquad \text{and} \qquad \lim_{\ell,q}\sum_{h>k} M(h,k) \,=\, 0.$$

Asymptotically, the coordinates $0,\dots,k$ of the normalized occupancy process $(O_n/m)_{n\geq 0}$ can be seen as a random perturbation of the discrete dynamical system given by the iterates of G:

$$r^0\in\mathcal{D}^k\,,\qquad r^n \,=\, G(r^{n-1}) \,=\, G^n(r^0)\,,\quad n\geq 1\,. \tag{DS_k}$$

In fact, this dynamical system will play a key role in our analysis. The mapping G and the dynamical system (DS_k) have extendedly been studied in the works [3] and [5]. The main results concerning the fixed points of G are given in Proposition 2.2 of [5], while the results concerning the stability of the fixed points and the convergence of the dynamical system are given in Theorems 3.1 and 4.1 in [5]. We summarize these results in the upcoming propositions. Consider the following set of indexes,

$$I_A \,:=\, \Big\{\, b\leq K \,\Big|\, A(b)e^{-a}>1 \text{ and } A(b)>A(j) \text{ for all } j>b \,\Big\} \cup \big\{\, K+1\,\big\}\,.$$

Proposition 2.5. *The mapping G has as many fixed points in $\mathcal{D}^k$ as there are elements in I_A. For each $b\in I_A$, the associated solution ρ^b is given by $\rho^b_0=\cdots=\rho^b_{b-1}=0$ and for $0\leq j\leq k-b$,*

$$\rho^b_{b+j} \,:=\, \Bigg(\frac{1}{A(b)}+\sum_{\substack{h\geq 1\\ 0=i_0<\cdots<i_h}}\frac{a^{i_h}}{A(b+i_h)}\prod_{t=1}^{h}\frac{A(b+i_t)}{(i_t-i_{t-1})!(A(b)-A(b+i_t))}\Bigg)^{-1}$$
$$\times\Bigg(\frac{1}{A(b)}1_{j=0}+\frac{a^j}{A(b+j)}\sum_{\substack{1\leq h\leq j\\ 0=i_0<\cdots<i_h=j}}\prod_{t=1}^{h}\frac{A(b+i_t)}{(i_t-i_{t-1})!(A(b)-A(b+i_t))}\Bigg)\,.$$

Note that in the case $b = K + 1$ the sum in the denominator diverges and therefore the solution ρ^{K+1} is identically 0. Moreover the trivial solution 0 is the only fixed point of G if and only if $A(0)e^{-a} \leq 1$. Even if we are defining the ρ^b only inside $\mathcal{D}^k$, we can actually consider the infinite sequence $(\rho^b_{b+k})_{k\geq 0}$; except for the trivial solution ρ^{K+1}, the other ρ^b all sum up to 1, and therefore give rise to probability distributions. Let $I_A = \{ b_1, \dots, b_N \}$ and note that $N = 1$ corresponds to 0 being the only fixed point of G. Define, for $b \in I_A$, the set $D_b \subset \mathcal{D}^k$ by

$$D_b \,:=\, \Big\{ r \in \mathcal{D}^k : r_0 + \cdots + r_{b-1} = 0 \Big\}\,.$$

We have the following result.

Proposition 2.6. *Let $r \in \mathcal{D}^k \setminus \{ 0 \}$. For every $i \in \{ 1, \dots N \}$,*

$$\lim_{n\to\infty} G^n(r) \,=\, \rho^{b_i}$$

if and only if

$$r_0 = \cdots = r_{b_{i-1}} = 0 \qquad \text{and} \qquad \max_{b_{i-1}<k\leq b_i} r_k > 0\,.$$

Moreover, the map G is contracting in a small enough neighborhood of ρ^b intersected with D_b.

Consider for example the following fitness function:

$$A(0) \,=\, 5\,, \quad A(1) \,=\, 2\,, \quad A(2) \,=\, 4\,, \quad A(3) \,=\, A(4) \,=\, \cdots \,=\, 1\,.$$

In this case $K = 2$. Suppose further that a is such that $4e^{-a} > 1 > 2e^{-a}$. Then, the mapping $G = (G_0, G_1, G_2)$ has three fixed points in the set $\mathcal{D}^2 = \{ r \in \mathbb{R}^3 : r_0, r_1, r_2 \geq 0$ and $r_0 + r_1 + r_2 \leq 1 \}$. The point 0 is always a fixed point, and in this case, its basin of attraction is just $\{ 0 \}$. We have two other fixed points, ρ^0 and ρ^2. The basin of attraction of ρ^2 is the set $\{ r \in \mathcal{D}^2 : r_0 = 0 \} \setminus \{ 0 \}$, and the basin of attraction of ρ^0 is the set $\{ r \in \mathcal{D}^2 : r_0 > 0 \}$. In fact, if $A(0)e^{-a} > 1$, the fixed point ρ^0 always exists, and its basin of attraction is always the set $\{ r \in \mathcal{D}^2 : r_0 > 0 \}$. Moreover, the mapping G is contracting in a small enough neighborhood of ρ^0.

2.6 Main result

Our main result concerns the invariant probability measure of the Wright–Fisher process. Recall the definition of the occupancy process $(O_n)_{n\geq 0}$ (Definition 2.2). The process $(O_n)_{n\geq 0}$ is recurrent and aperiodic, in fact, it can jump from any possible population to any other in just one step. Thus, it has a unique invariant probability measure. Let us denote by μ the invariant probability measure of the process $(O_n)_{n\geq 0}$. For any $0 \leq k \leq l$ we denote by π_k the mapping $\mathbb{R}^{l+1} \to \mathbb{R}^{k+1}$ that projects onto the the first $k + 1$ coordinates, i.e.

$$\forall\, x \in \mathbb{R}^{l+1} \qquad \pi_k(x) \,:=\, (x_0, \dots, x_k)\,. \tag{2.10}$$

For $k \geq 0$, let us denote by ν_k the image of the measure μ through the mapping $o \mapsto \pi_k(o/m)$. Our main result states that there is a dichotomy for the behavior of the measure ν_k. Depending on the values of the mean mutation rate per individual per reproduction cycle a and the rescaled population size α, the process will concentrate around the quasispecies fixed point ρ^0 or the disorder fixed point 0. Recall that the following limits are taken in the asymptotic regime (2.5). Recall also the definitions of A (Definition 2.1) and ρ^0 (Proposition 2.5).

Theorem 2.7. *There exists a function $\psi :]0,+\infty[\to [0,+\infty[$, which is finite on $]0,\ln A(0)[$ and vanishes on $[\ln A(0),+\infty[$, such that:*

- *If $\alpha\psi(a) < \ln\kappa$, then, for every $k \geq 0$, the measure ν_k converges weakly to the measure δ_0, i.e.,*

$$\lim_{\ell,m,q} \nu_k \to \delta_0 \, .$$

- *If $\alpha\psi(a) > \ln\kappa$, then, for every $k \geq 0$, the measure ν_k converges weakly to the measure $\delta_{(\rho_0^0,\dots,\rho_k^0)}$, i.e.,*

$$\lim_{\ell,m,q} \nu_k \to \delta_{(\rho_0^0,\dots,\rho_k^0)} \, .$$

In terms of the occupancy process $(O_n)_{n\geq 0}$, the above result can be restated as follows.

Corollary 2.8. *Let $(O_n)_{n\geq 0}$ be the occupancy process started from any configuration $o \in \mathcal{P}_{\ell+1}^m$. We have the following dichotomy, with the same function ψ as in Theorem 2.7:*

- *If $\alpha\psi(a) < \ln\kappa$ then*

$$\forall\, k \geq 0 \qquad \lim_{\ell,m,q} \lim_{n\to\infty} E\bigg(\frac{O_n(k)}{m}\bigg) \,=\, 0 \, .$$

- *If $\alpha\psi(a) > \ln\kappa$ then*

$$\forall\, k \geq 0 \qquad \lim_{\ell,m,q} \lim_{n\to\infty} E\bigg(\frac{O_n(k)}{m}\bigg) \,=\, \rho_k^0 \, .$$

Moreover, in both cases,

$$\forall\, k \geq 0 \qquad \lim_{\ell,m,q} \lim_{n\to\infty} Var\bigg(\frac{O_n(k)}{m}\bigg) \,=\, 0 \, .$$

Note that the phenomenon described by theorem 2.5 is a generalization of the corresponding dichotomy in the deterministic setting, and presents a richer behavior than its discrete counterpart. Indeed, as we can see from Theorem 3.3 in [5], whenever $A(0)e^{-a} > 1$, the unique fixed point of the mapping F converges to ρ^0, whereas when $A(0)e^{-a} \leq 1$, the unique fixed point of F converges to 0. In the deterministic setting, the population size is effectively infinite, whereas for the Wright–Fisher model, the extra parameter α corresponding to the population size comes into play. When $A(0)e^{-a} \leq 1$, the invariant measure of the Wright–Fisher model converges to δ_0, which is similar to the deterministic case. However, when $A(0)e^{-a} > 1$, the Wright–Fisher model possesses a richer behavior than its deterministic counterpart. When the population is below the threshold given by the curve $\alpha\phi(a) = \ln\kappa$, the invariant measure converges to δ_0, whereas when the population is above the threshold, the invariant measure converges to δ_{ρ^0}.

2.7 The phase transition

In the above theorem we can observe a phase transition given by the curve $\alpha\phi(a) = \ln\kappa$. The function ψ is defined as the solution to a variational problem, which in turn arises from a Large Deviations estimate for the occupancy process. For $p,t \in \mathcal{D}$, define the quantity $I_K(p,t)$ as follows (cf. equation (3.1)):

$$I_K(p,t) \,:=\, \sum_{k=0}^{K} t_k \ln\frac{t_k}{p_k} + (1-|t|_1)\ln\frac{1-|t|_1}{1-|p|_1} \, .$$

We make the convention that $0\ln 0 = 0\ln(0/0) = 0$. Define further $\rho(a)$ to be the quantity

$$\rho(a) \,:=\, \begin{cases} (\rho_0^0,\dots,\rho_K^0) & \text{if } A(0)e^{-a} > 1 \, , \\ (0,\dots,0) & \text{if } A(0)e^{-a} \leq 1 \, . \end{cases}$$

We can define the function ψ as

$$\psi(a) := \inf\left\{ \sum_{k=0}^{l} I_K\big(G(s^k), s^{k+1}\big) : \begin{array}{c} s^0=\rho(a), s^l=0, \\ s^k\in\mathcal{D}^K \text{ for } 0\le k\le l \end{array} \right\}. \tag{2.11}$$

We refer the reader to the Definition 9.1 and the comment thereafter for a more detailed explanation of this definition. Note that the dependence of the function ψ on a is through both the mapping G and the fixed point $\rho(a)$. When $A(0)e^{-a}\le 1$, we have $\rho(a)=0$, which is the unique stable fixed point of the mapping G, so that in this case, the infimum above is just 0. Proceeding as in Lemma 7.4 of [2], it ca be shown that when $A(0)e^{-a}>1$, the value $\psi(a)$ is strictly positive. In the case of the sharp peak landscape (i.e. $K=0$) we are able to exploit the monotonicity of the model and prove that the function $\psi(a)$ is decreasing in a, but unfortunately, right now we are not able to say the same for the general case.

2.8 Overview of the proof

Recall that the occupancy process can be seen as a random perturbation of the discrete–time dynamical system (DS_k). When $A(0)e^{-a}\le 1$, the mapping G has 0 as its only fixed point, and the result readily follows (we may write down the equation for the invariant measure of the Wright–Fisher model, and pass to the limit there, obtaining that the limiting measure must be invariant with respect to the action of G; the only such measure in the case $A(0)e^{-a}\le 1$ is the Dirac mass at 0). On the contrary, when $A(0)e^{-a}>1$, there are at least two fixed points, ρ^0 and 0, and the behavior of the process is much more intricate. Let us call these two points the *main* fixed points. We will refer to all other fixed points as the *intermediate* ones. The dichotomy presented in the theorem reflects a competition between the two main fixed points. The core of the proof of Theorem 2.7 lies in proving

$$\text{time}(\rho^0\to 0) \sim e^{m\psi(a)},$$
$$\text{time}(0\to\rho^0) \sim \kappa^\ell \sim e^{m\alpha^{-1}\ln\kappa},$$

along with showing that the time spent by the process away from the two main fixed points is negligible with respect to the above times. In the above formula the function ψ is the same as in Section 2.7. Recall that we have called the genotypes in the classes beyond K the *neutral* genotypes, and the ones in the classes $0,\dots,K$ the non–neutral ones. The behavior of the process is radically different according to whether there are *non–neutral* genotypes present in the population or not. Let us denote by $\mathcal{W}^*$ the set of populations containing at least one non–neutral individual, and by $\mathcal{N}$ the complementary set of populations, the all–neutral populations. The fixed point ρ^0, along with any existing intermediate fixed point, lies in the set $\mathcal{W}^*$, while the fixed point 0 lies in the set $\mathcal{N}$. In order to evaluate the above times, we carry out the same kind of estimations in both sets:

Time to the fixed point. We estimate the time that it takes to reach a neighborhood of the main fixed point (ρ^0 in $\mathcal{W}^*$ or 0 in $\mathcal{N}$), we develop this estimation uniformly on the starting population.

Time to leave the set. We estimate the time that the process needs to leave the set, and enter its complementary ($\mathcal{W}^*$ or $\mathcal{N}$). Once again, the estimation is developped uniformly on the strating population.

Excursions. We estimate the length and the frequency of the excursions outside a neighborhood of the main fixed point.

Even if the estimations we need are of the same nature for both sets and both main fixed points, the behavior of the process is very different in the neutral set and in the non–neutral one. As long as the process remains in the set $\mathcal{N}$, there is no selection, and the process behaves the same as if the function A were constant. In order to study this phase we will rely on the results found in [2] for the sharp–peak landscape, and we will show in Section 8 that the mean time needed to exit the neutral phase is of the order of κ^ℓ. The non–neutral phase consists of the populations where at least one of the classes $0,\dots,K$ is present. In the set of non–neutral populations, the process $(O_n)_{n\geq 0}$ will tend to behave as the dynamical system associated to G, this fact will be rigorously stated thanks to a large deviations principle, which we develop in Section 3. Inspired by the theory of Freidlin and Wentzell for random perturbations of dynamical systems [9], we exploit the large deviations principle in order to control several quantities associated with the process $(O_n)_{n\geq 0}$:

• We show that the process is very unlikely to stay away from a neighborhood of all of the fixed points for a long time (Section 4).

• We show that the process enters the basin of attraction of the main fixed point ρ^0 in a few steps with reasonable probability. In fact, this is one of the most technical parts of the proof, since the large deviations principle is of little help. Indeed, we need to control the probability for the process to create ηm master sequences out of 1 master sequence, for some $\eta > 0$ (Section 5).

• We estimate the mean time that the process needs to exit the set of non–neutral populations, which turns out to be of the order of $e^{m\psi(a)}$. The function $\psi(a)$ represents the quasipotential linking the points ρ^0 and 0, or otherwise stated, the "energy" of the most likely path the process is to follow when going from ρ^0 to 0 (Section 6).

• We show that when inside the set of non–neutral populations, the process spends most of its time in a neighborhood of ρ^0 (Section 7).

Finally, we put all the above estimates together, and we use them to show the main theorem, with help of the Ergodic Theorem for Markov chains (Section 9). The case $K = 0$ corresponds to the sharp–peak landscape, and has been treated in [2, 4]. The generalization to the class–dependent case is not straightforward. Indeed, the proofs in [2, 4] rely strongly on coupling and monotonicity arguments, which cease to work for arbitrary class–dependent functions. In addition, the behavior of the dynamical system associated to G is richer; in the sharp peak landscape, the only possible fixed points are ρ^0 and 0, while for more general fitness functions intermediate fixed points appear. When the occupancy process is in the set $\mathcal{N}$, the results in [2, 4] carry over to our case, however, when the occupancy process is in the set $\mathcal{W}^*$, this is no longer the case. To simplify the analysis when the process is in $\mathcal{W}^*$, we introduce a truncated process in the next subsection, and subsequently we develop new proofs for this truncated process. The new proofs rely on finding estimates that are uniform with respect to the initial points, and are therefore more robust than the original proofs in [2, 4].

2.9 Truncated process

Since our aim is to send the length of the sequences ℓ to infinity, the number of coordinates of the occupancy process will grow to infinity with ℓ. In order to deal with this inconvenience, we will truncate the process $(O_n)_{n\geq 0}$ so that the number of coordinates is fixed. Let k be an integer larger or equal to K. Recall the definitions of the process $(O_n)_{n\geq 0}$ (Definition 2.2) and of the mapping π_k (equation (2.10)). We define the truncated process $(Z_n)_{n\geq 0}$ by setting

$$\forall\, n \geq 0 \qquad Z_n := \pi_k\big(O_n\big)\,.$$

The process $(Z_n)_{n\geq 0}$ takes values in the set

$$\mathbb{D}^k := \left\{ z \in \mathbb{N}^{k+1} : z_0 + \cdots + z_k \leq m \right\}.$$

The process $(Z_n)_{n\geq 0}$ is not Markovian, since the coordinates that we are leaving out in its definition cannot be ignored when computing the transition probabilities of the process $(Z_n)_{n\geq 0}$. Indeed, for any $o \in \mathcal{P}^m_{\ell+1}$, $z \in \mathbb{D}^k$ and $n \geq 0$ we have

$$P\big(Z_{n+1} = z \,\big|\, O_n = o\big) = \frac{m!}{z_1! \cdots z_k!(m - |z|_1)!} \times F_0\Big(\frac{o}{m}\Big)^{z_0} \cdots F_k\Big(\frac{o}{m}\Big)^{z_k} \left(1 - \left|\pi_k\Big(F\Big(\frac{o}{m}\Big)\Big)\right|_1\right)^{m-|z|_1}. \quad (2.12)$$

However, in the asymptotic regime we consider, the process $(Z_n)_{n\geq 0}$ behaves as a small random perturbation of the dynamical system associated to the mapping G, and therefore, the process $(Z_n)_{n\geq 0}$ can be seen as being “asymptotically Markovian”. There is an exception to this, which is when the occupancy process in in the set $\mathcal{N}$ (c.f. the Section 2.8). However, as long as the process remains in $\mathcal{N}$, the occupancy process $(O_n)_{n\geq 0}$ behaves as if the evolution was neutral ($A \equiv 0$). This situation has been thoroughly studied in Section 8 of [2], and most of the results therein carry over to our case, as we discuss in the Section 8. In the remaining subsequent sections the process $(Z_n)_{n\geq 0}$ will be the main object of our study. We develop next a Large Deviations Principle for the transition probabilities of the process $(Z_n)_{n\geq 0}$.

2.10 Notation

In defining $(Z_n)_{n\geq 0}$ we have fixed a coordinate $k \geq K$, but since the treatment is the same for all $k \geq K$, in the sequel, we assume that $k = K$.

- For $k \geq K$ we define the set $\mathcal{D}^k := \left\{ r \in [0,1]^{k+1} : |r|_1 \leq 1 \right\}$.
- For a subset B of $\mathcal{D}^k$, we denote by $\mathbb{B}$ the set $mB \cap \mathcal{D}$.
- We denote the sets $\mathbb{D}^K$ and $\mathcal{D}^K$ simply by $\mathbb{D}$ and $\mathcal{D}$.
- We denote the mapping $\pi_K(\cdot)$ by $\pi(\cdot)$.
- $\mathcal{D}_\gamma$ denotes the set $\{\, r \in \mathcal{D} : r_0 \geq \gamma \,\}$, and $\mathbb{D}_\gamma$ denotes its discrete counterpart $m\mathcal{D}_\gamma \cap \mathbb{D}$.
- For $r \in \mathbb{R}^d$, we denote by $|r|_1$ the 1–norm of r.
- For $r \in \mathbb{R}^d$, we denote by $\lfloor r \rfloor$ the vector $\lfloor r \rfloor = (\lfloor r_0 \rfloor, \ldots, \lfloor r_d \rfloor)$.
- For $x \in \mathcal{D}$ and $\delta > 0$ we denote $U_\delta(x)$ or $U(x, \delta)$ the δ–neighborhood of x in $\mathcal{D}$.
- For $b \in I_A$ we use the shorthand $U^b_\delta = U_\delta(\rho^b)$.
- We write U_δ for $\cup_{b\in I_A} U^b_\delta$, and W_δ for $U_\delta \setminus (U_\delta(\rho^0) \cup U_\delta(0))$.
- We write $\overline{\phi}(B)$ and $\underline{\phi}(B)$ for the maximum and minimum fitness over $r \in B \subset \mathcal{D}$. We use the shorthand $\overline{\phi}^b_\delta := \overline{\phi}(U^b_\delta)$ and $\underline{\phi}^b_\delta := \underline{\phi}(U^b_\delta)$.
- For any set $B \in \mathbb{D}$ we write $\tau(B)$ for the hitting time of B by the process $(Z_n)_{n\geq 0}$.

In the sequel, by *asymptotically* we mean: for ℓ, m large enough, q small enough, ℓq close enough to a and m/ℓ close enough to α. All subsequent statements and inequalities need not be true for all values of ℓ, m and q, but only asymptotically, even if we do not state so explicitly. For $o \in \mathcal{P}^m_{\ell+1}$ or $z \in \mathbb{D}$, we use the notation

$$E_o(\cdot) := E(\cdot \,|\, O_0 = o), \qquad E_z(\cdot) := E_z(\cdot \,|\, Z_0 = z).$$

Note that in general the state of the process Z_1 given $Z_0 = z$ is not well defined, it depends on the missing coordinates too, thus, expressions of the sort

$$P_z(Z_1 = 0) \geq p\,,$$

should be interpreted as

$$P_o(Z_1 = 0)\ \geq p \text{ for all } o \text{ such that } \pi(o) = z\,.$$

3 Large deviations principle

In the asymptotic regime (2.5), the trajectories of the process $(Z_n)_{n\geq 0}$ (cf. the Section 2.9 for its definition) follow closely the trajectories of the dynamical system (DS_k). In order to quantify what "follows closely" means, we develop here a Large Deviations Principle for the process $(Z_n)_{n\geq 0}$. For $p,t \in \mathcal{D}$, we define the quantity $I_K(p,t)$ as follows:

$$I_K(p,t)\ :=\ \sum_{k=0}^{K} t_k \ln\frac{t_k}{p_k} + (1-|t|_1)\ln\frac{1-|t|_1}{1-|p|_1}\,, \tag{3.1}$$

We make the convention that $0\ln 0 = 0\ln(0/0) = 0$. The function $I_K(p,\cdot)$ is the rate function governing the large deviations of a multinomial distribution (see for example [11]) with parameters n and $p_0,\dots,p_K, 1-|p|_1$. We have the following estimate for the multinomial coefficients:

Lemma 3.1. *Let $n \geq N \geq 1$, and $i_1,\dots,i_N \in \mathbb{N}$ be such that $i_1+\cdots+i_N = n$. We have*

$$\left|\ln\frac{n!}{i_1!\cdots i_N!} + \sum_{k=1}^{N} i_k \ln\frac{i_k}{n}\right|\ \leq\ N\ln n + 2N\,.$$

The proof is similar to that of Lemma 7.1 of [2]. Taking logarithms in (2.12), we obtain

$$\begin{aligned}\ln\, P_o\big(Z_{n+1}=z\big)\ &=\ \ln\,\frac{m!}{z_1!\cdots z_n!(m-|z|_1!)} + \sum_{0\leq i\leq k} z_i \ln F_i\Big(\frac{o}{m}\Big)\\ &+ (m-|z|_1)\ln\bigg(1-\Big|\pi\big(F(o/m)\big)\Big|_1\bigg)\ =\ \ln\,\frac{m!}{z_1!\cdots z_n!(m-|z|_1!)}\\ &- mI_K\Big(\pi\big(F(o/m)\big), z/m\Big) + \sum_{0\leq i\leq k} z_i\ln z_i + \big(m-|z|_1\big)\ln\big(m-|z|_1\big)\,.\end{aligned}$$

Thanks to the lemma, for $o \in \mathcal{P}^m_{\ell+1}$ and $z \in \mathbb{D}$

$$\ln P_o\big(Z_{n+1}=z\big)\ =\ -mI_K\Big(\pi\big(F(o/m)\big), z/m\Big) + \Phi(o,z)\,. \tag{3.2}$$

The error term $\Phi(o,z)$ satisfies, for m large enough,

$$\forall\, o\in\mathcal{P}^m_{\ell+1}\quad \forall\, z\in\mathbb{D}\qquad \big|\Phi(o,z)\big|\ \leq\ C(K)\ln m\,, \tag{3.3}$$

where $C(K)$ is a constant that depends on K but not on m. We define a function $V_1 : \mathcal{D}\times\mathcal{D} \to [0,\infty]$ by setting, for $r,t \in \mathcal{D}$,

$$V_1(r,t)\ :=\ I_K\big(G(r),t\big)\,. \tag{3.4}$$

Since the F_i converge to the G_i (cf. (2.9)), we have,

$$\forall x\in\mathcal{S}^\ell\quad \forall t\in\mathcal{D}\qquad \lim_{\ell,q} I_K\Big(\pi\big(F(x)\big),t\Big)\ =\ V_1(\pi(x),t)\,.$$

Proposition 3.2. *The one step transition probabilities of the Markov chain $(Z_n)_{n\geq 0}$ satisfy the large deviations principle governed by V_1:*

- *For any subset B of $\mathcal{D}$ and for any $r\in\mathcal{D}$, we have, for $n\geq 0$,*

$$-\inf\big\{\,V_1(r,t):t\in\overset{o}{B}\,\big\}\,\leq\,\liminf_{\ell,m,q}\frac{1}{m}\ln P\big(Z_{n+1}\in\mathbb{B}\,\big|\,Z_n=\lfloor mr\rfloor\big)\,.$$

- *For any subsets B,B' of $\mathcal{D}$, we have, for $n\geq 0$,*

$$\limsup_{\ell,m,q}\frac{1}{m}\ln\sup_{z\in\mathbb{B}}P_z\big(Z_{n+1}\in\mathbb{B}'\big)\leq\,-\inf\big\{\,V_1(r,t):r\in\overline{B},\,t\in\overline{B}'\,\big\}\,.$$

This result is similar to the well known Gärtner–Ellis Theorem. Nevertheless, in our case we have three parameters instead of one, and a series of measures that depend on the the states $z\in\mathbb{D}$ and $r\in\mathcal{D}$. In order to deal with this particularities, we deem it necessary to prove this proposition from the scratch.

Proof. We begin by showing the large deviations upper bound. Let B,B' be two subsets of $\mathcal{D}$ and notice that, for all $z\in\mathbb{D}$ and $n\geq 0$

$$P_z\big(Z_{n+1}\in\mathbb{B}'\big)\,\leq\,\sup_{o:\pi(o)=z}P_o\big(Z_{n+1}\in\mathbb{B}'\big)\,.$$

Let $o\in\mathcal{P}^m_{\ell+1}$ be such that $\pi(o)\in\mathbb{B}$. For $n\geq 0$, we have

$$P_o\big(Z_{n+1}\in\mathbb{B}'\big)\,=\,\sum_{z'\in\mathbb{B}'}P_o\big(Z_{n+1}=z'\big)\,.$$

The number of elements in the sum is of polynomial order in m, the exponent of m depending on K only. Indeed, $|\mathbb{B}'|\leq|\mathbb{D}|<m^{K+1}$. Thus, thanks to the estimates (3.2) and (3.3) on the transition probabilities for the process $(Z_n)_{n\geq 0}$, we have, for m large enough,

$$\begin{aligned}\sup_{o:\pi(o)\in\mathbb{B}}P_o\big(Z_{n+1}\in\mathbb{B}'\big)\,&\leq\,m^{C(K)}\sup_{\substack{o:\pi(o)\in\mathbb{B}\\ z'\in\mathbb{B}'}}P_o\big(Z_{n+1}=z'\big)\\ &\leq\,m^{C'(K)}\exp\bigg(-m\min_{\substack{o:\pi(o)\in\mathbb{B}\\ z'\in\mathbb{B}'}}I_K\Big(\pi\big(F(o/m)\big),z'/m\Big)\bigg)\,.\end{aligned}$$

where $C(K)$ and $C'(K)$ are constants that depend on K but not on m. Define the mappings $\underline{F},\overline{F}:\mathcal{D}\longrightarrow\mathcal{D}$ by setting, for all $r\in\mathcal{D}$ and $k\in\{\,0,\dots,K\,\}$

$$\underline{F}_k(r):=\frac{1}{\phi(r)}\sum_{0\leq i\leq k}r_iA(i)M(i,k),\tag{3.5}$$

$$\overline{F}_k(r):=\frac{1}{\phi(r)}\bigg(\sum_{0\leq i\leq k}r_iA(i)M(i,k)+\big(1-|\pi_k(r)|_1\big)A(0)M(k+1,k)\bigg).\tag{3.6}$$

Asymptotically, for $0\leq k<j\leq\ell$, we have $M(j,k)\leq M(k+1,k)$ (cf. the Subsection 2.4). Thus, asymptotically, for all x in the unit simplex $\mathcal{S}^\ell$, and for all $k\in\{\,0,\dots,K\,\}$,

$$\underline{F}_k\big(\pi(x)\big)\,\leq\,F_k(x)\,\leq\,\overline{F}_k\big(\pi(x)\big)\,.$$

Define next the function $\underline{V}:\mathcal{D}\times\mathcal{D}\longrightarrow[0,+\infty]$ by

$$\forall\,r,t\in\mathcal{D}\qquad\underline{V}(r,t)\,:=\,\sum_{i=0}^K t_i\ln\frac{t_i}{\overline{F}_i(r)}+\big(1-|t|_1\big)\ln\frac{1-|t|_1}{1-\big|\pi(\underline{F}(r))\big|_1}\,.$$

Asymptotically, the function $\underline{V}$ satisfies

$$\forall\, x\in\mathcal{S}^\ell\quad\forall\, t\in\mathcal{D}\qquad \underline{V}\big(\pi(x),t\big)\ \le\ I_K\big(\pi(F(x)),t\big)\,.$$

Moreover, asymptotically, for $r,t\in\mathcal{D}$, we have $\underline{V}(r,t)\ \longrightarrow\ V_1(r,t)$. Thus,

$$\sup_{o:\pi(o)\in\mathbb{B}} P_o\big(Z_{n+1}\in\mathbb{B}'\big)\ \le\ (m+1)^{C'(K)}\exp\bigg(-m\min_{z\in\mathbb{B},z'\in\mathbb{B}'}\underline{V}\Big(\frac{z}{m},\frac{z'}{m}\Big)\bigg)\,.$$

For each $m\ge 1$, let $z_m,z'_m\in\mathbb{D}$, be two terms that realize the above minimum. Up to the extraction of a subsequence, we can suppose that when $m\to\infty$,

$$\frac{z_m}{m}\to r\in\overline{B}\,,\qquad \frac{z'_m}{m}\to t\in\overline{B}'\,.$$

Thus,

$$\limsup_{\ell,m,q} -\underline{V}\Big(\frac{z_m}{m},\frac{z_m}{m}\Big)\ \le\ -V_1(r,t)\,.$$

Optimizing with respect to r,t, we obtain the upper bound of the large deviations principle. Let $r,t\in\mathcal{D}$ and notice that, for all $z\in\mathbb{D}$ and $n\ge 0$

$$P_z\big(Z_{n+1}=\lfloor mt\rfloor\big)\ \ge\ \inf_{o:\pi(o)=z} P_o\big(Z_{n+1}=\lfloor mt\rfloor\big)\,.$$

let $o\in\mathcal{P}^m_{\ell+1}$ be such that $\pi(o)=\lfloor mr\rfloor$. We have

$$P_o\big(Z_{n+1}=\lfloor mt\rfloor\big)\ge\ m^{-C(K)}\exp\bigg(-mI_K\Big(\pi\big(F(o/m)\big),\lfloor mt\rfloor/m\Big)\bigg),$$

where $C(K)$ is a constant depending on K but not on m. Define the function $\overline{V}:\mathcal{D}\times\mathcal{D}\longrightarrow[0,+\infty]$ by

$$\forall\, r,t\in\mathcal{D}\qquad \overline{V}(r,t)\ :=\ \sum_{i=0}^K t_i\ln\frac{t_i}{\underline{F}_i(r)}+\big(1-|t|_1\big)\ln\frac{1-|t|_1}{1-\big|\pi(\overline{F}(r))\big|_1}\,.$$

Asymptotically, the function $\overline{V}$ satisfies

$$\forall\, x\in\mathcal{S}^\ell\quad\forall\, t\in\mathcal{D}\qquad \overline{V}\big(\pi(x),t\big)\ \ge\ I_K\big(\pi(F(x)),t\big)\,.$$

Moreover, asymptotically, for $r,t\in\mathcal{D}$,

$$\overline{V}(r,t)\ \longrightarrow\ V_1(r,t)\,.$$

Thus, for every $o\in\mathcal{P}^m_{\ell+1}$ such that $\pi(o)\ =\ \lfloor mr\rfloor$,

$$P_o\big(Z_{n+1}=\lfloor mt\rfloor\big)\ \ge\ m^{-C(K)}\exp\bigg(-m\overline{V}\Big(\frac{\lfloor mr\rfloor}{m},\frac{\lfloor mt\rfloor}{m}\Big)\bigg)\,.$$

We take the logarithm and we send m,ℓ to ∞ and q to 0. We obtain then

$$\liminf_{\ell,m,q}\frac{1}{m}\ln P\big(Z_{n+1}=\lfloor tm\rfloor\,\big|\,Z_n=\lfloor rm\rfloor\big)\ \ge\ -V_1(r,t)\,.$$

Moreover, if $t\in\overset{o}{U}$, for m large enough, $\lfloor tm\rfloor$ belongs to $\mathbb{B}$. Therefore,

$$\liminf_{\ell,m,q}\frac{1}{m}\ln P\big(Z_{n+1}\in\mathbb{B}\,\big|\,Z_n=\lfloor mr\rfloor\big)\ \ge\ -V_1(r,t)\,.$$

We optimize over t and we obtain the large deviations lower bound. □

The l–step transition probabilities of $(Z_n)_{n\geq 0}$ also satisfy a large deviations principle. For $l \geq 2$, we define a function V_l on $\mathcal{D}\times\mathcal{D}$ as follows:

$$V_l(r,t) := \inf\Big\{ \sum_{k=0}^{l-1} V_1(s^k,s^{k+1}) : s^0 = r,\ s^l = t,\ s^k \in \mathcal{D} \text{ for } 0 \leq k \leq l \Big\}. \tag{3.7}$$

Corollary 3.3. *For $l \geq 1$, the l–step transition probabilities of $(Z_n)_{n\geq 0}$ satisfy the large deviations principle governed by V_l:*

- *For any subset B of $\mathcal{D}$ and for any $r \in \mathcal{D}$, we have, for $n \geq 0$,*

$$-\inf\big\{ V_l(r,t) : t \in \overset{o}{B} \big\} \leq \liminf_{\ell,m,q} \frac{1}{m} \ln P\big(Z_{n+l} \in \mathbb{B} \,\big|\, Z_n = \lfloor mr \rfloor\big).$$

- *For any subsets B, B' of $\mathcal{D}$, we have, for $n \geq 0$,*

$$\limsup_{\ell,m,q} \frac{1}{m} \ln \sup_{z\in\mathbb{B}} P\big(Z_{n+l} \in \mathbb{B}' \,\big|\, Z_n = z\big) \leq -\inf\big\{ V_l(r,t) : r \in \overline{B},\ t \in \overline{B}' \big\}.$$

We omit the proof of the corollary, since it is similar to that of the Theorem 3.2, combined with an induction argument based on the Markov property. The rate function $V_1(r,t)$ is equal to 0 if and only if $t = G(r)$. Thus, the Markov chain $(Z_n/m)_{n\geq 0}$ can be seen as a random perturbation of the dynamical system associated to the map G (cf. Section 2.1). The next sections study the consequences of the large deviations principles of Proposition 3.2 and Corollary 3.3 on the asymptotic behavior of the process $(Z_n)_{n\geq 0}$.

4 Time spent away from the fixed points

The aim of this section is to show that the process $(Z_n)_{n\geq 0}$ has a small probability of staying away from a neighborhood of the fixed points for a long time. We begin by giving a useful lemma. We have discussed the behavior of the dynamical system (DS_k) in Section 2.5, we recall that the set U_δ is a union of δ–neighborhoods of the fixed points (cf. Section 2.10). The set $\mathcal{D} \setminus U_\delta$ is compact and for every $r \in \mathcal{D} \setminus U_\delta$,

$$\lim_{n\to\infty} G^n(r) \in U_\delta.$$

Lemma 4.1. *Let $\delta > 0$. There exist $h \in \mathbb{N}$ and $c > 0$ (depending on δ) such that, asymptotically, for every point $z \in \mathbb{D} \setminus \mathbb{U}_\delta$*

$$P_z\Big(Z_1 \notin \mathbb{U}_\delta, \dots, Z_h \notin \mathbb{U}_\delta\Big) \leq e^{-cm}.$$

Proof. Recall that for $r \in \mathcal{D}$ we denote by r^n the n–th iterate of r by the map G. By continuity of the map G, for every $r \in \mathcal{D}\setminus U_\delta$, there exists $h(r) \in \mathbb{N}$ and $0 < \eta_0^r, \dots, \eta_{h(r)}^r < \varepsilon$ such that, for all $0 \leq n \leq h(r) - 1$,

$$G\big(U(r^n,\eta_n^r)\big) \subset U(r^{n+1},\eta_{n+1}^r/2) \quad \text{and} \qquad U(r^{h(r)},\eta_{h(r)}^r) \subset U_\delta \tag{4.1}$$

The family $\{ U(r,\eta_0^r) : r \in \mathcal{D} \setminus U_\delta \}$ forms an open cover of the compact $\mathcal{D} \setminus U_\delta$. Thus, there exist $r_1,\dots,r_M \in \mathcal{D} \setminus U_\delta$ such that

$$\mathcal{D} \setminus U_\delta \subset \bigcup_{1\leq i\leq M} U(r_i,\eta_0^{r_i}).$$

Set

$$h := \max_{1\leq i\leq M} h(r_i).$$

Let $t \in \mathcal{D} \setminus U_\delta$ and let $i \in \{ 1, \dots, M \}$ be such that $t \in U(r_i, \eta_0^{r_i})$. We denote the quantity $h(r_i)$ simply by $h(i)$, the open ball $U(r_i^n, \eta_n^{r_i})$ by U_n, and we set $\mathbb{U}_n = mU_n \cap \mathbb{D}$. We have then,

$$\begin{aligned} P_z\big(Z_1 \notin \mathbb{U}_\delta, \dots, Z_h \notin \mathbb{U}_\delta\big) &\leq P_z\big(Z_n \notin \mathbb{U}_n \text{ for some } 1 \leq n \leq h(i)\big) \\ &\leq \sum_{1 \leq n \leq h(i)} P_z\big(Z_1 \in \mathbb{U}_1, \dots, Z_{n-1} \in \mathbb{U}_{n-1}, Z_n \notin \mathbb{U}_n\big) \\ &\leq \sum_{1 \leq n \leq h(i)} \sum_{z' \in \mathbb{U}_{n-1}} P_z\big(Z_n \notin \mathbb{U}_n \,\big|\, Z_{n-1} = z'\big) P_z\big(Z_{n-1} = z'\big) . \end{aligned}$$

The large deviations principle for the transitions of $(Z_n)_{n \geq 0}$ yields the following bound,

$$\limsup_{\ell, m, q} \frac{1}{m} \ln P_{z'}\big(Z_n \notin \mathbb{U}_n\big) \leq -\inf \big\{ V_1(\rho, \rho') : \rho \in U_{n-1}, \rho' \notin U_n \big\} = -c_i^n .$$

Since $G(U_{n-1}) \subset U_n$, the constant c_i^n is strictly positive. The number of constants $(c_i^n)_{1 \leq n \leq h(i), 1 \leq i \leq M}$ is finite. Let $0 < \eta < \min_{i,n} c_i^n$. From the above inequalities, we conclude that, asymptotically,

$$\begin{aligned} &P_{\lfloor tm \rfloor}\big(Z_1 \notin \mathbb{U}_\delta, \dots, Z_h \notin \mathbb{U}_\delta\big) \\ &\quad \leq \sum_{1 \leq n \leq h(i)} \exp\big(-m(c_i^n - \eta)\big) P_z\big(Z_{n-1} \in \mathbb{U}_{n-1}\big) \leq \prod_{1 \leq n \leq h(i)} \exp\big(-m(c_i^n - \eta)\big) . \end{aligned}$$

Since $h(i) \leq h$ and h is fixed, and since the number of constants c_i^n is finite, the above probability is bounded by e^{-mc}, for $c > 0$ independent of t. □

We use the lemma to prove the following corollary.

Corollary 4.2. *There exist $h \in \mathbb{N}$ and $c > 0$ such that, asymptotically, for every $z \in \mathbb{D} \setminus \mathbb{U}_\delta$ and $n \in \mathbb{N}$,*

$$P_z\Big(Z_t \notin \mathbb{U}_\delta, 0 \leq t \leq n\Big) \leq \exp\Big(-mc\Big\lfloor \frac{n}{h} \Big\rfloor\Big) .$$

Proof. Divide the interval $\{ 0, \dots, n \}$ into subintervals of length h. Using iteratively the previous lemma, we have, for $i \geq 1$,

$$\begin{aligned} P_z\big(Z_t \notin \mathbb{U}_\delta, 0 \leq t \leq (i+1)h\big) &= \sum_{z' \in \mathbb{D} \setminus \mathbb{U}_\delta} P_z\big(Z_t \notin \mathbb{U}_\delta, 0 \leq t \leq (i+1)h, Z_{ih} = z'\big) \\ &= \sum_{z' \in \mathbb{D} \setminus \mathbb{U}_\delta} P_z\big(Z_t \notin \mathbb{U}_\delta, 0 \leq t \leq ih, Z_{ih} = z'\big) \\ &\quad \times P_z\big(Z_t \notin \mathbb{U}_\delta, ih < t \leq (i+1)h \big| Z_{ih} = z'\big) \leq P_z\big(Z_t \notin \mathbb{U}_\delta, 0 \leq t \leq ih\big) e^{-mc}. \end{aligned}$$

Iterating this procedure we get

$$P_z\big(Z_t \notin \mathbb{U}_\delta, 0 \leq t \leq (i+1)h\big) \leq e^{-mc(i+1)} .$$

Taking $i + 1 = \lfloor n/h \rfloor$ gives the desired result. □

Recall the definitions of $\mathcal{D}_\gamma$, U_δ^0 and $\tau(\cdot)$ from Section 2.10. Let $\gamma, \delta > 0$. In view of Proposition 2.6, for any $r \in \mathcal{D}_\gamma$, the trajectory $G^n(r)$ converges to ρ^0, and does so without getting close to 0. We have results analogous to Lemma 4.1 and Corollary 4.2 for this case.

Lemma 4.3. *Let $\delta, \gamma > 0$,. There exist $h \in \mathbb{N}$ and $c > 0$ (depending on δ, γ) such that, asymptotically, for every point $z \in \mathbb{D}_\gamma \setminus \mathbb{U}_\delta^0$*

$$P_z\Big(\tau(\mathbb{U}_\delta^0 \cup \{0\}) \le h,\ Z_{\tau(\mathbb{U}_\delta^0 \cup \{0\})} \neq 0\Big) \ \ge\ 1 - e^{-cm}\,.$$

Corollary 4.4. *There exist $h \in \mathbb{N}$ and $c > 0$ such that, asymptotically, for every $z \in \mathbb{D}_\gamma \setminus \mathbb{U}_\delta^0$ and $n \in \mathbb{N}$,*

$$P_z\Big(\tau(\mathbb{U}_\delta^0 \cup \{0\}) \le n,\ Z_{\tau(\mathbb{U}_\delta^0 \cup \{0\})} \neq 0\Big) \ \ge\ 1 - \exp\Big(-mc\Big\lfloor \frac{n}{h}\Big\rfloor\Big)\,.$$

The proofs of these two results are straightforward modifications of the proofs of Lemma 4.1 and Corollary 4.2, so we omit them.

5 Creating enough master sequences

Throughout this whole section we assume that $A(0)e^{-a} > 1$. The aim of this section is to show that starting from any point of $\mathbb{D} \setminus \{\,0\,\}$, the process $(Z_n)_{n \ge 0}$ creates a number of master sequences of order m with a reasonable probability, within a time of order $\ln m$.

Theorem 5.1. *Let $\epsilon > 0$. There exist positive constants γ and C such that for every $z \in \mathbb{D} \setminus \{\,0\,\}$,*

$$\liminf_{\ell,m,q} \frac{1}{m} \ln P_z\big(\tau(\mathbb{D}_\gamma \cup \{0\}) \le C \ln m,\ Z_{\tau(\mathbb{D}_\gamma \cup \{0\})} \neq 0\big) \ \ge\ -\epsilon\,.$$

Assume first that the process $(Z_t)_{t \ge 0}$ starts from a neighborhood of one of the fixed points. More precisely, let $b \in I_A \setminus \{\,0\,\}$ and assume that the starting point is in a small neighborhood of ρ^b and is of the form

$$z \ =\ (w, 0, \dots, 0, z_b, \dots, z_K)\,.$$

Since, for δ small enough, G is contracting in the intersection of the set $D_b = \{\, r \in \mathcal{D} : r_0 + \cdots + r_{b-1} = 0\,\}$ with a sufficiently small neighborhood of ρ^b, the process $(Z_t)_{t \ge 0}$ will stay inside such a neighborhood for a long time. Note that for some $\varepsilon > 0$ depending on the neighborhood,

$$G_0\Big(\frac{z}{m}\Big) \ \ge\ \frac{wA(0)e^{-a}}{A(b)e^{-a} + \varepsilon}\,.$$

A similar inequality holds for points close to z, so that if the neighborhood is small enough, as long as the process is inside it, the number of master sequences will tend to increase geometrically. This is the key idea of the proof of the theorem, which will be carried out in a few different steps:

• First we show that from any starting point, the process jumps to a point of the form of z in a finite number of steps, with probability higher than $e^{-\varepsilon m}$, for every $\varepsilon > 0$.

• Then we build a deterministic trajectory that, starting from a point of the form of z, creates γm master sequences in a time of order $\ln m$.

• Finally we show that the process is likely enough to follow the deterministic trajectory.

This strategy will be implemented in a series of lemmas, which we state next. The proofs of the lemmas will be carried out after we give the proof of the Theorem 5.1. Recall the definitions of U_δ^b, U_δ and $\tau(\cdot)$ from Section 2.10.

Lemma 5.2. *Let $\delta > 0$. There exist $h \in \mathbb{N}$ and $c > 0$ such that, asymptotically, for every $z \in \mathbb{D} \setminus \{0\}$,*

$$P_z\big(\tau(\mathbb{U}_\delta) \le h,\ Z_{\tau(\mathbb{U}_\delta)} \neq 0\big) \ \ge\ 1 - e^{-cm}\,.$$

We omit the proof of this lemma due to its similarity with Lemma 4.1. Indeed, the only difference between this lemma and Lemma 4.1, is the fact that here we don't want the chain to hit 0. Since $G(z) = 0$ if and only if $z = 0$, it is enough to modify slightly the proof of Lemma 4.1 (by replacing $\mathbb{U}_\delta$ by $\mathbb{U}_\delta \setminus \{0\}$ in (4.1) and where subsequently needed).

Lemma 5.3. *Let $\varepsilon > 0$ and $w \in \mathbb{N}$. Taking δ small enough, for any $b \in I_A$ and $z \in \mathbb{U}_\delta^b \setminus \{0\}$,*

$$\lim_{\ell,q,m} \frac{1}{m} \ln P_z\big(Z_1 = (w, 0, \dots, 0, z_b, \dots, z_K)\big) \;\geq\; -\epsilon\,.$$

Suppose next that

$$z \;=\; (w, 0, \dots, 0, z_b, \dots, z_K) \;\in\; \mathbb{U}_{\delta/2}^b\,.$$

We build a deterministic trajectory $(z^n)_{n\geq 0}$ such that $z^0 = z$ and for δ small enough, w large enough, the number of master sequences grows geometrically. In order to do so we use the auxiliary mapping $\underline{F}$ defined in (3.5). The deterministic trajectory is built by setting $z^0 = z$ and for $n \geq 1$,

$$z^n \;=\; \left\lfloor m\underline{F}\Big(\frac{z^{n-1}}{m}\Big)\right\rfloor.$$

Recall the definition of ϕ_δ^b from Section 2.10. Take δ small enough and w large enough so that, asymptotically,

$$\rho \;:=\; \frac{A(0)M(0,0)}{\overline{\phi}} - \frac{1}{w} \;>1\,.$$

Note that letting δ go to zero, ℓ to infinity, q to 0 and ℓq to a, the quantity ρ converges to $A(0)/A(b) - 1/w$. So that asymptotically, for δ below a certain threshold δ_0 and w above a certain threshold w_0, the quantity ρ is always bounded below by some quantity $\overline{\rho} > 1$. Moreover, the conditions $\delta < \delta_0$ and $w > w_0$ are not contradictory, since as long as $w \leq \delta m/2$, the point z will remain inside $\mathbb{U}_\delta^b$. We define by N_δ^b the first time of exit of the dynamical system z^n from $\mathbb{U}_\delta^b$:

$$N_\delta^b \;:=\; \inf\big\{\, n \geq 0 : z^n \notin \mathbb{U}_\delta^b \,\big\}\,.$$

Lemma 5.4. *The sequence $(z_n^0)_{0\leq n\leq N_\delta^b}$ is increasing, and bounded below by a geometric sequence with ratio ρ.*

Proof. Note first that

$$z_0^1 \;=\; \left\lfloor \frac{wA(0)M(0,0)}{\phi(z^0/m)} \right\rfloor \;\geq\; w\bigg(\frac{A(0)M(0,0)}{\overline{\phi}_\delta^b} - \frac{1}{w}\bigg) \;=\; \rho w\,.$$

Then, by induction, for any $n \leq N_\delta^b$,

$$z_0^n \;=\; \left\lfloor \frac{z_0^{n-1}A(0)M(0,0)}{\phi(z^{n-1}/m)} \right\rfloor \;\geq\; z_0^{n-1}\bigg(\frac{A(0)M(0,0)}{\overline{\phi}_\delta^b} - \frac{1}{w}\bigg) \;=\; \rho z_0^{n-1}\,. \qquad \square$$

Let $\gamma > 0$, then, $\rho^n w$ is larger than γm if

$$n \;\geq\; n(\gamma) \;:=\; \big(\ln\rho\big)^{-1} \ln\frac{\gamma m}{w}\,.$$

Note that as mentioned before ρ is bounded below by $\overline{\rho} > 1$, and that w is of order one while m is going to infinity. So that for any fixed $\gamma > 0$, the quantity $n(\gamma)$ also goes to infinity with m. The deterministic trajectory needs $n(\gamma)$ steps to create γm master sequences. But the bound $z_0^n \geq \rho^n w$ only works as long as the trajectory remains inside the set $\mathbb{U}_\delta^b$, therefore, we need to ensure that the trajectory has the time to take $n(\gamma)$ steps before exiting $\mathbb{U}_\delta^b$.

Lemma 5.5. *Let $\delta > 0$ and $b \in I_A$. For γ small enough, $n(\gamma) \leq N_\delta^b$.*

Finally, we need a lemma ensuring that our process is likely to follow the deterministic trajectory we have just built.

Lemma 5.6. *Let $\delta > 0$, $z^0 = (w, 0, \dots, z_b, \dots, z_K) \in \mathbb{U}^b_{\delta/2}$ and let $(z^n)_{n\geq 0}$ be the trajectory built from z^0 by setting $z^n = \lfloor m\underline{F}(z^{n-1}/m)\rfloor$. Then, for γ small enough, we have,*

$$\liminf_{\ell,q} \frac{1}{m} \ln P_{z^0}\big(Z_{n(\gamma)} = z^{n(\gamma)}, \dots, Z_1 = z^1\big) \,=\, 0\,.$$

Let us show next how to combine the previous lemmas in order to prove Theorem 5.1.

Proof of Theorem 5.1. Let $\varepsilon > 0$ as in Lemma 5.3. Take $\delta, \gamma > 0$ and $w \in \mathbb{N}$ so that they meet the requirements of Lemmas 5.3, 5.4 and 5.5. We define the following events:

$$\begin{aligned}
\mathcal{E}_1(\delta) \,&:=\, \Big\{\, \tau(\mathbb{U}_\delta) \leq h,\; Z_{\tau(\mathbb{U}_\delta)} \neq 0 \,\Big\}\,,\\
\mathcal{E}_2(\delta, w) \,&:=\, \mathcal{E}_1 \cap \Big\{\, Z_{\tau(\mathbb{U}_\delta)+1} = (w, 0, \dots, 0, z_b, \dots, z_K) \text{ for some b} \in I_A \,\Big\}\,,\\
\mathcal{E}_3(\delta, w, \gamma) \,&:=\, \mathcal{E}_2 \cap \Big\{\, Z_{\tau(\mathbb{U}_\delta)+2} = z^1, \dots, Z_{\tau(\mathbb{U}_\delta)+1+n(\gamma)} = z^{n(\gamma)} \,\Big\}\,.
\end{aligned}$$

Note that if $\mathcal{E}_3$ is realized, we have $Z_{\tau(\mathbb{U}_\delta)+1+n(\gamma)} \geq \gamma m$. Let C be such that

$$C \ln m \,\geq\, h + 1 + n(\gamma)\,.$$

Then,

$$P_z\big(\tau(\mathbb{D}_\gamma \cup \{0\}) \leq C \ln m\big) \,\geq\, P_z\big(\mathcal{E}_3 \,\big|\, \mathcal{E}_2\big) P_z\big(\mathcal{E}_2 \,\big|\, \mathcal{E}_1\big) P_z\big(\mathcal{E}_1\big)\,.$$

Now, by Lemma 5.2, for any $\delta > 0$

$$\liminf_{\ell,m,q} \frac{1}{m} \ln P_z\big(\mathcal{E}_1\big) \,=\, 0\,.$$

Moreover, by Lemma 5.3, for any $\varepsilon > 0$ and $w \in \mathbb{N}$, we can choose δ small enough so that

$$\liminf_{\ell,m,q} \frac{1}{m} \ln P_z\big(\mathcal{E}_2 \,\big|\, \mathcal{E}_1\big) \,\geq\, -\epsilon\,.$$

Finally, by Lemma 5.6, choosing γ small enough,

$$\liminf_{\ell,m,q} \frac{1}{m} \ln P_z\big(\mathcal{E}_3 \,\big|\, \mathcal{E}_2\big) \,=\, 0\,.$$

Combining these three limits we get the desired result. ☐

Once the process Z_n has γm master sequences, it will converge to ρ^0 in a few steps, as shown in the following corollary.

Corollary 5.7. *Let $\varepsilon, \delta > 0$. There exists a positive constant C such that for every $z \in \mathbb{D} \setminus \{\,0\,\}$*

$$\liminf_{\ell,m,q} \frac{1}{m} \ln P_z\big(\tau(\mathbb{U}^0_\delta \cup \{0\}) \leq C \ln m,\; Z_{\tau(\mathbb{U}^0_\delta \cup \{0\})} \neq 0\big) \,\geq\, -\varepsilon\,.$$

Proof. Let $\varepsilon, \delta > 0$, and let $\gamma, C' > 0$ associated to ε as in Theorem 5.1. Then, for any $C > C'$ and $z \in \mathbb{D} \setminus \{\,0\,\}$, we have,

$$\begin{aligned}
P_z\begin{pmatrix} \tau(\mathbb{U}^0_\delta \cup \{0\}) \leq C \ln m, \\ Z_{\tau(\mathbb{U}^0_\delta \cup \{0\})} \neq 0 \end{pmatrix} \,\geq\, &\sum_{z' \in \mathbb{D} : z'(0) \geq \gamma m} P_z\begin{pmatrix} Z_{\tau(\mathbb{D}_\gamma \cup \{0\})} = z', \\ \tau(\mathbb{D}_\gamma \cup \{0\}) < C' \ln m \end{pmatrix}\\
&\qquad \times P_z\left(\begin{matrix} \tau(\mathbb{U}^0_\delta \cup \{0\}) \leq C \ln m, \\ Z_{\tau(\mathbb{U}^0_\delta \cup \{0\})} \neq 0 \end{matrix} \,\middle|\, \begin{matrix} Z_{\tau(\mathbb{D}_\gamma \cup \{0\})} = z, \\ \tau(\mathbb{D}_\gamma \cup \{0\}) < C' \ln m \end{matrix}\right).
\end{aligned}$$

By Lemma 4.3, there exist $h \in \mathbb{N}$ and $c > 0$ such that, asymptotically, for every $z' \in \mathbb{D}$ satisfying $z'_0 \geq \gamma m$

$$P_{z'}\big(\tau(\mathbb{U}^0_\delta \cup \{0\}) \leq h,\ Z_{\tau(\mathbb{U}^0_\delta \cup \{0\})} \neq 0\big) \,\geq\, 1 - e^{-cm}\,.$$

Thus, taking C such that $\lfloor C \ln m \rfloor - \lfloor C' \ln m \rfloor > h$, and in view of Theorem 5.1, we obtain the result of the corollary. ☐

It remains to prove Lemmas 5.3, 5.5 and 5.6. Before we do so, let us introduce some preliminary notation and results. We will make use of the mappings $\underline{F}, \overline{F}$ defined in (3.5) and (3.6). Recall that the mappings $\underline{F}$ and $\overline{F}$ satisfy, asymptotically,

$$\forall\, x \in \mathcal{S}^\ell \quad \forall\, k \in \{\,0, \dots, K\,\} \qquad \underline{F}_k\big(\pi(x)\big) \,\leq\, F_k(x) \,\leq\, \overline{F}_k\big(\pi(x)\big)\,.$$

Proposition 5.8. *We have,*

$$\lim_{\ell,q} \sup_{r \in \mathcal{D}} \big|\underline{F}(r) - G(r)\big|_1 \,=\, 0\,, \qquad \lim_{\ell,q} \sup_{r \in \mathcal{D}} \big|\overline{F}(r) - G(r)\big|_1 \,=\, 0\,.$$

Proof. Recall that M_∞ represents the limit mutation matrix (cf. the Subsection 2.4). Let $r \in \mathcal{D}$ and $k \in \{\,0, \dots, K\,\}$, we have

$$\big|G_k(r) - \underline{F}_k(r)\big| \,\leq\, \phi(r)^{-1} \sum_{0 \leq i \leq k} r_i A(i) \big|M_\infty(i,k) - M(i,k)\big| \,\leq$$
$$\sup_{r \in \mathcal{D}} \bigg(\phi(r)^{-1} \sum_{0 \leq i \leq k} r_i A(i)\bigg) \max_{0 \leq i,k \leq K} \big|M_\infty(i,k) - M(i,k)\big|\,.$$

This last quantity converges to 0 asymptotically, uniformly on $r \in \mathcal{D}$. The rest of the lemma can be shown in a similar way. ☐

Results similar to Propositions 2.5 and 2.6 hold for the mapping $\underline{F}$. The proofs are exactly the same, even if the form of the fixed points is different. More precisely, we have the following result.

Proposition 5.9. *Asymptotically, the mapping $\underline{F}$ has as many fixed points in $\mathcal{D}$ as there are elements in I_A. For each $b \in I_A$, the associated fixed point η^b satisfies $\eta^b_0 = \cdots = \eta^b_{b-1} = 0$ and $\eta^b_b \wedge \cdots \wedge \eta^b_K > 0$. The mapping $\underline{F}$ restricted to $\mathcal{D}_b$ is contracting in a neighborhood of η^b, and*

$$\lim_{\ell,q} \eta^b \,=\, \rho^b\,.$$

The last convergence is a direct consequence of Proposition 5.8. Let $\varepsilon > 0$. The mapping G is continuous on the compact set $\mathcal{D}$, it is therefore uniformly continuous on $\mathcal{D}$. In view of the Proposition 5.8, δ can be chosen small enough so that for all $b \in I_A$, asymptotically,

$$\big|\overline{F}(r) - \eta^b\big| + \big|\underline{F}(r) - \eta^b\big| < \varepsilon\,.$$

Proof of Lemma 5.3. Let $\varepsilon > 0$. Suppose now that $z \in \mathbb{U}^b_\delta$ for some $b \in I_A$, fix $w \in \mathbb{N}$ and set $z' = (w, 0, \dots, 0, z_b, \dots, z_K)$. We wish to show that, for δ small enough,

$$\lim_{\ell,q} \frac{1}{m} \ln P_z\big(Z_1 = z'\big) \,\geq\, -\varepsilon\,.$$

Note that for any $o \in \mathcal{P}^m_{\ell+1}$ satisfying $\pi(o) = z \in \mathbb{D} \setminus \{0\}$, we have the following asymptotic bound on the probability of creating a master sequence,

$$F_0\Big(\frac{o}{m}\Big) \,\geq\, \phi\Big(\frac{z}{m}\Big)^{-1} \sum_{0 \leq h \leq K} \frac{z_h}{m} A(h) M(h,0) \,\geq\, m^{-M}\,,$$

for some $M > 0$, which is uniform over $z \in \mathbb{D} \setminus \{0\}$. This bound comes from the fact that $M(h,0)$ is of order $q^h \sim m^{-h}$. Thus, we obtain the following lower bound on the probability of jumping from z to z',

$$\frac{m!}{w! z'_b! \cdots z'_K! (m - |z'|_1)!} m^{-Mw} \prod_{i=b}^{K} \underline{F}_i\Big(\frac{z}{m}\Big)^{z'_i} \times \bigg(1 - \Big|\overline{F}\Big(\frac{z}{m}\Big)\Big|_1\bigg)^{m-|z'|_1}.$$

We use now the Lemma 3.1 in order to obtain the following asymptotic bound,

$$\begin{aligned} \frac{1}{m} \ln P\big(Z_1 = z' \,\big|\, Z_0 = z\big) \;\geq\; & -\frac{1}{m}(K+1)(2+\ln m) - \frac{w}{m} \ln \frac{w/m}{m^{-M}} \\ & - \sum_{i=b}^{K} \frac{z'_i}{m} \ln \frac{z'_i/m}{\underline{F}_i(z/m)} - \frac{m - |z'|_1}{m} \ln \frac{(m - |z'|_1)/m}{1 - |\overline{F}(z/m)|_1}. \end{aligned}$$

The first two quantities go to 0, when m goes to infinity, so that the sum of both is eventually larger that $-\varepsilon/2$. Since $z/m \in U^b_\delta$, we have

$$\Big|\frac{z}{m} - \underline{F}\Big(\frac{z}{m}\Big)\Big|_1 \;\leq\; \Big|\frac{z}{m} - \eta^b\Big| + \Big|\eta^b - \underline{F}\Big(\frac{z}{m}\Big)\Big| \;<\; \delta + \varepsilon\,.$$

Furthermore, for $b \leq i \leq K$, the function $\underline{F}_i(r)$ is bounded below by a positive constant c when $r \in U^b_\delta$, we conclude that

$$\sum_{i=b}^{K} \frac{z_i}{m} \ln \frac{z_i/m}{\underline{F}_i(z/m)} \;\leq\; K \ln\Big(1 + \frac{\delta+\varepsilon}{c}\Big)\,.$$

We choose δ small enough so that this last quantity is smaller than $\varepsilon/4$. A similar argument shows that δ can be chosen small enough so that the last term is also bounded below by $-\varepsilon/4$, thus giving the desired bound. □

In order to prove Lemma 5.5, it is enough to ensure that while the number of master sequences is growing, the rest of the coordinates don't change too fast, so that the trajectory can remain inside $\mathbb{U}^b_\delta$ for more than $n(\gamma)$ steps. In order to do so, we show in the next lemma that the coordinates $(z^n_k)_{0\leq k<b}$ cannot grow at a faster rate than z^n_0. We prove afterwards that the same thing holds for the difference $|z^n_k - \eta^b_k|$, where $b \leq k \leq K$. Let us choose $\varepsilon > 0$ small enough so that, asymptotically,

$$\forall\, k \in \{\,1,\dots,K\,\} \qquad \frac{A(k)M(k,k)}{A(0)M(0,0) - \varepsilon} \;<\; 1\,,$$

and let w be such that $\overline{\phi}^b_\delta/w < \varepsilon$, where $\overline{\phi}^b_\delta$ denotes the maximum fitness in U^b_δ (cf. the Section 2.10). We prove next the following lemma.

Lemma 5.10. *There exist positive constants $c_0,\dots,c_{b-1}$ such that for $0 \leq k < b$, asymptotically, $z^n_k \leq c_k z^n_0$, for all $n \leq N^b_\delta$.*

Proof. We will prove the lemma by induction on k. The case $k = 0$ is obviously true. Set $k \in \{\,1,\dots,b-1\,\}$ and suppose that the statement of the theorem holds for the coordinates $0,\dots,k-1$. Then, we have

$$\frac{z^n_k}{z^n_0} \;\leq\; \frac{\phi(z^{n-1}/m)^{-1} \displaystyle\sum_{0\leq i\leq k} z^{n-1}_i A(i)M(i,k)}{\phi(z^{n-1}/m)^{-1} z^{n-1}_0 A(0)M(0,0) - 1} \;\leq\; \sum_{0\leq i\leq k} \frac{z^{n-1}_i A(i)M(i,k)}{z^{n-1}_0 A(0)M(0,0) - \overline{\phi}^b_\delta}\,.$$

The sequence $(z_0^n)_{0\leq n\leq N_\delta^b}$ is increasing, and thus, $\overline{\phi}_\delta^b/z_0^{n-1} < \overline{\phi}_\delta^b/w < \varepsilon$. Therefore,

$$\frac{z_k^n}{z_0^n} \;\leq\; \sum_{0\leq i\leq k} \frac{z_i^{n-1}}{z_0^{n-1}}\left(\frac{A(i)M(i,k)}{A(0)M(0,0)-\varepsilon}\right).$$

By the induction hypothesis

$$\frac{z_k^n}{z_0^n} \;\leq\; \sum_{0\leq i\leq k-1} c_i\left(\frac{A(i)M(i,k)}{A(0)M(0,0)-\varepsilon}\right) + \frac{z_k^{n-1}}{z_0^{n-1}}\frac{A(k)M(k,k)}{A(0)M(0,0)-\varepsilon}\,.$$

Iterating this inequality, and noting that $z_k^0=0$, we obtain

$$\frac{z_k^n}{z_0^n} \;\leq\; \sum_{0\leq i\leq k-1} c_i\left(\frac{A(i)M(i,k)}{A(0)M(0,0)-\varepsilon}\right)\sum_{t=0}^{n-1}\left(\frac{A(k)M(k,k)}{A(0)M(0,0)-\varepsilon}\right)^t.$$

Yet, asymptotically,

$$d_k \;:=\; \frac{A(k)M(k,k)}{A(0)M(0,0)-\varepsilon} \;<\; 1\,,$$

and thus,

$$\frac{z_k^n}{z_0^n} \;\leq\; \frac{1}{1-d_k}\sum_{0\leq i\leq k-1} c_i\left(\frac{A(i)M(i,k)}{A(0)M(0,0)-\varepsilon}\right).$$

We take c_k to be equal to the right hand side of this inequality, which fulfills the proof of the lemma. □

Let us introduce the following notation, for any $x\in\mathbb{R}^{K+1}$ we denote by $\widetilde{x}$ the vector x where the coordinates $0,\dots,b-1$ have been set to 0, i.e.,

$$\widetilde{x} \;:=\; (0,\dots,0,x_b,\dots,x_K)\,.$$

Lemma 5.11. *There exist constants $c_b>0$ and $0<c_\delta<1$ such that for $n\leq N_\delta^b$, asymptotically, we have*

$$\big|\widetilde{z^n}-m\eta^b\big|_1 \;\leq\; c_b z_0^n + c_\delta^n\big|\widetilde{z^0}-m\eta^b\big|_1\,.$$

Proof. For $n\geq 0$ and $b\leq k\leq K$, we have, noting that η^b is a fixed point of the mapping $\underline{F}$,

$$\big|z_k^n-m\eta_k^b\big| \;\leq\; 1+m\big|\underline{F}_k(z^{n-1}/m)-\underline{F}_k(\eta^b)\big|\,.$$

Yet, for all $r\in\mathcal{D}$

$$\underline{F}_k(r) \;=\; \sum_{i=0}^{b-1} r_i\frac{A(i)M(i,k)}{\phi(r)} + \underline{F}_k(\widetilde{r})\frac{\phi(\widetilde{r})}{\phi(r)}\,.$$

Thus,

$$\begin{aligned}\big|z_k^n-m\eta_k^b\big| \;\leq\; 1+\sum_{i=0}^{b-1} z_i^{n-1}\frac{A(i)M(i,k)}{\phi(z^{n-1}/m)} + m\big|\underline{F}_k(\widetilde{z}^{n-1}/m)-\underline{F}_k(\eta^b)\big| \\ + \underline{F}_k(\widetilde{z}^{n-1}/m)\frac{|\phi(\widetilde{z^{n-1}}/m)-\phi(z^{n-1}/m)|}{\phi(z^{n-1}/m)}\,.\end{aligned}$$

We have,

$$\big|\phi(\widetilde{z^{n-1}}/m)-\phi(z^{n-1}/m)\big| \;\leq\; \sum_{j=0}^{b-1}\frac{z_j^{n-1}}{m}|A(j)-1|$$

Reporting back in the inequality for $|z_k^n - m\eta_k^b|$, we get,

$$\big|z_k^n - m\eta_k^b\big| \;\leq\; 1 + m\Big|\underline{F}_k(\widetilde{z^{n-1}}/m) - \underline{F}_k(\eta^b)\Big| \\ + \sum_{j=0}^{b-1} z_j^{n-1}\left(\frac{A(j)M(j,k) + |A(j)-1|\underline{F}_k(\widetilde{z^{n-1}}/m)}{\phi(z^{n-1}/m)}\right).$$

Summing from $k=b$ to K, and recalling that, asymptotically, $\underline{F}$ is contracting on the set $U_\delta^b \cap \mathcal{D}_b$, we deduce the existence of a constant $c_\delta < 1$ such that

$$\big|\widetilde{z^n} - m\eta^b\big|_1 \leq K + c_\delta\big|\widetilde{z^{n-1}} - m\eta^b\big|_1 + \sum_{j=0}^{b-1} z_j^{n-1}\frac{A(j)\displaystyle\sum_{b\leq k\leq K} M(j,k) + |A(j)-1|}{\overline{\phi}_\delta^b}.$$

Using the previous lemma, we get that, asymptotically,

$$\big|\widetilde{z^n} - m\eta^b\big|_1 \;\leq\; C z_0^n + c_\delta\big|\widetilde{z^{n-1}} - m\eta^b\big|_1\,,$$

for some constant C depending on δ only. Iterating this inequality, and noting that for $t \leq n$, we have $z_0^{n-t} \leq \rho^{-t} z_0^n$, we conclude that

$$\big|\widetilde{z^n} - m\eta^b\big|_1 \;\leq\; C\sum_{t=0}^{n-1} c_\delta^t z_0^{n-t} + c_\delta^n\big|\widetilde{z^0} - m\eta^b\big|_1 \\ \leq\; C z_0^n \sum_{t=0}^{n-1}(c_\delta/\rho)^t + c_\delta^n\big|\widetilde{z^0} - m\eta^b\big|_1 \;\leq\; C\big(1 - c_\delta/\rho\big)^{-1} z_0^n + c_\delta^n\big|\widetilde{z^0} - m\eta^b\big|_1\,.$$

The proof is concluded by taking $c_b = C(1 - c_\delta/\rho)^{-1}$. □

Proof of Lemma 5.5. As a consequence of Lemmas 5.10 and 5.11, as long as $n \leq N_\delta^b$, we have

$$z_0^n \;\geq\; \frac{z_1^n}{c_1} \vee \cdots \vee \frac{z_{b-1}^n}{c_{b-1}} \vee \frac{|\widetilde{z^n} - m\eta^b|_1 - |\widetilde{z^0} - m\eta^b|_1}{c_b}\,.$$

Thus, taking $\gamma(K+1) < \min\{\,\delta/c_1, \dots, \delta/2c_b\,\}$, then $N_\delta^b \geq n(\gamma)$, as wanted. □

Proof of Lemma 5.6. We have,

$$P_{z^0}\big(Z_{n(\gamma)} = z^{n(\gamma)}, \dots, Z_1 = z^1\big) \;=\; \prod_{n=0}^{n(\gamma)-1} P_{z^0}\big(Z_{n+1} = z^{n+1}\,\big|\, Z_n = z^n\big)\,.$$

For any $0 \leq n \leq n(\gamma)$, as in the proof of the Large Deviations Principle 3.2,

$$P_{z^0}\big(Z_{n+1} = z^{n+1}\,\big|\, Z_n = z^n\big) \;\geq\; \inf_{o:\pi(0)=z^n} P\big(Z_{n+1} = z^{n+1}\,\big|\, O_n = o\big)\,.$$

Let $o \in \mathcal{P}_{\ell+1}^m$ be such that $\pi(o) = z^n$. We have,

$$P\big(Z_{n+1} = z^{n+1}\,\big|\, O_n = o\big) \;=\; \frac{m!}{z_0^{n+1}!\cdots z_K^{n+1}!(m - |z^{n+1}|_1)!} \\ \times F_0\Big(\frac{o}{m}\Big)^{z_0^{n+1}} \cdots F_K\Big(\frac{o}{m}\Big)^{z_K^{n+1}}\left(1 - \left|\pi\Big(F\Big(\frac{o}{m}\Big)\Big)\right|\right)^{m-|z^{n+1}|_1}.$$

Thus,

$$\ln P\big(Z_{n+1}=z^{n+1}\,\big|\,O_n=o\big)\;=\;-mI_K\bigg(\pi\Big(F\Big(\frac{o}{m}\Big)\Big),\frac{z^{n+1}}{m}\bigg)+\Phi(o,z^{n+1})\,,$$

where the error term $\Phi(o,z^{n+1})$ satisfies

$$\big|\Phi(o,z^{n+1})\big|\;\leq\;C(K)(\ln m+1)\,,$$

$C(K)$ being a constant that depends on K but not on m (c.f. the formulas (3.2) and (3.3)). Next we bound the quantity involving the rate function I. Recall that

$$mI_K\bigg(\pi\Big(F\Big(\frac{o}{m}\Big)\Big),\frac{z^{n+1}}{m}\bigg)$$
$$=\sum_{k=0}^{K}z_k^{n+1}\ln\frac{z_k^{n+1}/m}{F_k(o/m)}+\big(m-|z^{n+1}|_1\big)\ln\frac{1-|z^{n+1}|_1/m}{1-|\pi(F(o/m))|_1}\,.$$

The function $\underline{F}$ has been defined so that for all $x\in\mathcal{S}^\ell$, $\underline{F}_k\big(\pi(x)\big)\leq F_k(x)$, for $0\leq k\leq K$. Therefore, for all $o\in\mathcal{P}^m_{\ell+1}$ such that $\pi(o)=z^n$,

$$\frac{z_k^{n+1}}{m}\;=\;\frac{1}{m}\Big\lfloor m\underline{F}_k\Big(\frac{z^n}{m}\Big)\Big\rfloor\;\leq\;\underline{F}_k\Big(\frac{z^n}{m}\Big)\;\leq\;F_k\Big(\frac{o}{m}\Big)\,.$$

Thus,

$$mI_K\bigg(\pi\Big(F\Big(\frac{o}{m}\Big)\Big),\frac{z^{n+1}}{m}\bigg)\;\leq\;\big(m-|z^{n+1}|_1\big)\ln\frac{1-|z^{n+1}|_1/m}{1-|\pi(F(o/m))|_1}\,.$$

The argument of the logarithm is larger than 1, and for all $x\geq 0$, $\ln(x)\leq x-1$. Therefore, the above quantity is bounded by

$$\big(m-|z^{n+1}|_1\big)\frac{\big|\pi\big(F(o/m)\big)\big|_1-|z^{n+1}|_1/m}{1-\big|\pi\big(F(o/m)\big)\big|_1}$$
$$\leq\;\frac{m-|z^{n+1}|_1}{1-\big|\pi\big(F(o/m)\big)\big|_1}\sum_{k=0}^{K}\bigg|\frac{1}{m}\Big\lfloor m\underline{F}_k\Big(\frac{z^n}{m}\Big)\Big\rfloor-F_k\Big(\frac{o}{m}\Big)\bigg|$$
$$\leq\;\frac{m-|z^{n+1}|_1}{1-\big|\pi\big(F(o/m)\big)\big|_1}\sum_{k=0}^{K}\bigg(\Big|\underline{F}_k\Big(\frac{z^n}{m}\Big)-F_k\Big(\frac{o}{m}\Big)\Big|+\frac{1}{m}\bigg)\,.$$

For any $x\in\mathcal{S}^\ell$

$$\big|\pi\big(F(x)\big)\big|_1\;=\;\phi\big(\pi(x)\big)^{-1}\sum_{h=0}^{\ell}x_hA(h)\sum_{k=0}^{K}M(h,k)\,.$$

On one hand, for any $h\in\{\,0,\dots,\ell\,\}$ the sum $M(h,0)+\cdots+M(h,K)$ is bounded by a constant c which is strictly smaller than 1. Thus, $|\pi(F(x))|_1$ is bounded above by this same constant c, uniformly on $x\in\mathcal{S}^\ell$. Therefore, $(m-|z^{n+1}|_1)/(1-|\pi(F(o/m))|_1)\leq m/(1-c)$. On the other hand, for $0\leq k\leq K$, since $\pi(o)=z^n$,

$$\bigg|\underline{F}_k\Big(\frac{z^n}{m}\Big)-F_k\Big(\frac{o}{m}\Big)\bigg|\;=\;\phi(z^n/m)^{-1}\sum_{h=k+1}^{\ell}\frac{o_h}{m}A(h)M(h,k)\,.$$

Yet, there exists a positive constant c' such that asymptotically, $M(h,k)\leq c'/m$, for $0\leq k<h\leq\ell$. Therefore, the above quantity is bounded by c'/m. We conclude that

$$I_K\bigg(\pi\Big(F\Big(\frac{o}{m}\Big)\Big),\frac{z^{n+1}}{m}\bigg)\;\leq\;\frac{(1+c')(K+1)}{m(1-c)}\,.$$

Therefore,

$$\frac{1}{m}\ln P_{z^0}\big(Z_{n(\gamma)}=z^{n(\gamma)},\dots,Z_1=z^1\big)$$
$$\geq\ -\sum_{n=0}^{n(\gamma)-1}\inf_{o:\pi(o)=z^n}\left(I_K\left(\pi\Big(F\Big(\frac{o}{m}\Big)\Big),\frac{z^{n+1}}{m}\right)+\frac{1}{m}\Phi(o,z^{n+1})\right)$$
$$\geq -\frac{n(\gamma)(1+c')(K+1)}{m(1-c)}-\frac{n(\gamma)C(K)(\ln m+1)}{m}\,.$$

Since $n(\gamma)$ is of the order of $\ln m$, we see that this last quantity goes to 0 when m goes to infinity, as wanted. □

6 Persistence time

We assume throughout this whole section that $A(0)e^{-a}>1$. The aim of this section is to compute the expected hitting time of 0 for the process $(Z_n)_{n\geq 0}$. The relevant quantity for the computation is the quasipotential

$$V(r,t)\ :=\ \inf_{l\geq 1}V_l(r,t)\,, \tag{6.1}$$

where the quantity V_l has been defined in (3.7). Let us denote $\tau_0=\tau(\{0\})$ the hitting time of 0.

Theorem 6.1. *For all $z\in\mathbb{D}\setminus\{0\}$,*

$$\lim_{\ell,m,q}\frac{1}{m}\ln E_z(\tau_0)\ =\ V(\rho^0,0)\,.$$

In order to ease the readability of the upcoming formulas, we denote the quantity $V(\rho^0,0)$ simply as V.

Proof. We begin by showing the upper bound: for any $\varepsilon>0$, we have

$$\forall\, z\in\mathbb{D}\setminus\{\,0\,\}\qquad \limsup_{\ell,m,q}\frac{1}{m}\ln E_z(\tau_0)\ \leq\ V+\varepsilon\,.$$

Let $\varepsilon>0$. We first show that there exists a constant $C>0$ such that

$$\forall\, z\in\mathbb{D}\setminus\{\,0\,\}\qquad P_z\big(\tau_0\leq\lfloor C\ln m\rfloor\big)\ \geq\ e^{-m(V+2\varepsilon)}$$

Let $\gamma>0$, $z\in\mathbb{D}\setminus\{\,0\,\}$, and assume first that $z_0>\gamma m$. Define the sequence $(r^n)_{n\geq 0}$ by setting $r^0=z/m$ and

$$r^n\ =\ G^n(r^0)\,,\qquad n\geq 1\,.$$

The mapping V_1 is continuous on the first argument in a neighborhood of ρ^0; let us choose δ small enough so that

$$|r-\rho^0|_1\ <\ \delta\quad\Rightarrow\quad V_1(r,\rho^0)\ <\ \varepsilon/3\,. \tag{6.2}$$

Moreover, for δ small enough there exists $h\in\mathbb{N}$ such that for all $r\in\mathcal{D}_\gamma$, and for all $n\geq h$, we have

$$\big|G^n(r)-\rho^0\big|_1\ <\ \delta\,.$$

Indeed, by the Proposition 2.6, δ can be chosen sufficiently small so that the δ–neighborhood of ρ^0 is contracting. By continuity of the map G, for all $r\in\mathcal{D}_\gamma$, there exists $\delta_r>0$ and $h(r)\in\mathbb{N}$ such that if

$$|r-t|<\delta_r\quad\Longrightarrow\quad\big|G^{h(r)}(t)-\rho^0\big|\ <\ \delta\,.$$

The set $\mathbb{D}_\gamma$ is compact and the family $\{\,U(r,\delta_r) : r \in \mathbb{D}_\gamma\,\}$ is an open cover of the set $\mathbb{D}_\gamma$. Thus, there exist $r_1,\dots,r_N \in \mathbb{D}_\gamma$ such that

$$\mathbb{D}_\gamma \,\subset\, \bigcup_{i=1}^{N} U(r_i,\delta_{r_i})\,.$$

Set h to be the maximum of $h(r_1),\dots,h(r_N)$. Then, for all $r \in \mathbb{D}_\gamma$ and $n \geq h$, we have $|G^n(r)-\rho^0| < \delta$. Let $h' \geq 0$ and let $(t^i)_{0\leq i\leq h'}$ be a sequence in $\mathcal{D}$ satisfying

$$t^0 \,=\, \rho^0\,,\; t^{h'} \,=\, 0\,,\qquad \sum_{i=0}^{h'-1} V_1(t^i,t^{i+1}) \,\leq\, V+\frac{\varepsilon}{3}\,. \tag{6.3}$$

Consider next the sequence $(s_i)_{0\leq i\leq h+h'+1}$ defined by

$$s_0 = r^0\,,\quad s_1 = r^1\,,\quad \dots\,,\quad s_h = r^h\,,$$
$$s_{h+1} = t^0 = \rho^0\,,\quad s_{h+2} = t^1\,,\quad \dots\,,\quad s_{h+h'+1} = t^{h'} = 0\,.$$

Set $L = h+h'+1$. Combining (6.2) and (6.3) we see that the sequence $(s_i)_{0\leq i\leq L}$ satisfies

$$\sum_{i=0}^{L-1} V_1(s_i,s_{i+1}) \,\leq\, V+2\varepsilon/3\,.$$

Proceeding as in the proof of the Large Deviations Principle 3.2, we obtain

$$P_z\big(Z_N=0\big) \,\geq\, \prod_{t=0}^{L-1} P_z\big(Z_{t+1}=\lfloor s_{t+1}m\rfloor \,\big|\, Z_t=\lfloor s_t m\rfloor\big)\,.$$

Then,

$$\liminf_{\ell,m,q} \frac{1}{m}\,\ln P_z\big(Z_L=0\big) \,\geq\, -V+2\varepsilon/3\,.$$

Thus, asymptotically,

$$P_z\big(Z_L=0\big) \,\geq\, e^{-m(V+\varepsilon)}\,,$$

uniformly on $z \in \mathbb{D}_\gamma$. Suppose now that $z \notin \mathbb{D}_\gamma$. By Theorem 5.1, there exist $C' > 0$ such that

$$\forall\, z \in \mathbb{D}\setminus\{\,0\,\}\qquad P_z\big(Z_{\lfloor C'\ln m\rfloor} \geq \gamma m\big) \,\geq e^{-\varepsilon m}\,.$$

Thus, for every $z \in \mathbb{D}\setminus\{\,0\,\}$,

$$\begin{aligned} P_z\big(Z_{\lfloor C'\ln m\rfloor+L}=0\big) \,&\geq\, \sum_{z'\in\mathbb{D}_\gamma} P_z\big(Z_{\lfloor C'\ln m\rfloor+L}=0, Z_{\lfloor C'\ln m\rfloor}=z'\big)\\ &\geq\, \sum_{z'\in\mathbb{D}_\gamma} P_{z'}\big(Z_L=0\big)P_z\big(Z_{\lfloor C'\ln m\rfloor}=z'\big) \,\geq\, e^{-m(V+2\varepsilon)}\,. \end{aligned}$$

Taking C such that $\lfloor C\ln m\rfloor \geq \lfloor C'\ln m\rfloor + L$, we conclude that for every $z \in \mathbb{D}$,

$$P_z\big(\tau_0 \leq \lfloor C\ln m\rfloor\big) \,\geq\, e^{-m(V+2\varepsilon)}\,. \tag{6.4}$$

Proceeding as in Corollary 4.2 we obtain that, for every $h \geq 1$ and $z \in \mathbb{D}\setminus\{\,0\,\}$,

$$P_z\big(\tau_0 \geq h\lfloor C\ln m\rfloor\big) \,\leq\, \big(1-e^{-m(V+2\varepsilon)}\big)^h\,.$$

Thus,

$$E_z(\tau_0) \;=\; \sum_{n\geq 0} P_z\big(\tau_0 \geq n\big) \;=\; \sum_{h\geq 0} \sum_{n=h\lfloor C\ln m\rfloor}^{(h+1)\lfloor C\ln m\rfloor -1} P_z\big(\tau_0\geq n\big)$$
$$\leq \sum_{h\geq 0} \lfloor C\ln m\rfloor P_z\big(\tau_0 \geq h\lfloor C\ln m\rfloor\big) \;\leq\; \lfloor C\ln m\rfloor e^{m(V+2\varepsilon)}\,.$$

We conclude that, for every $z\in\mathbb{D}$,

$$\limsup_{\ell,m,q} \frac{1}{m}\ln E_z(\tau_0) \;\leq\; V+2\varepsilon\,.$$

We send ε to 0 and we obtain the desired upper bound. We proceed now to the proof of the lower bound. We will make use of the following inequality: let $\varepsilon>0$., for all $T>e^{2\varepsilon m}$ and for all $z\in\mathbb{D}\setminus\{\,0\,\}$, we have, asymptotically,

$$P_z\big(Z_t\notin \mathbb{U}_\delta(\rho^0)\cup\mathbb{U}_\delta(0), 0\leq t\leq T\big) \;\leq\; e^{-\varepsilon m}\,. \tag{6.5}$$

In order to prove this inequality we use the estimate of Corollary 5.7 and proceed exactly as in the proof of Corollary 4.2. Due to the similarity with the proof of Corollary 4.2, we omit the proof of the inequality. In order to prove the lower bound, we define τ_δ to be the hitting time of the δ–neighborhood of 0,

$$\tau_\delta \;:=\; \inf\big\{\,n\geq 0 : Z_n\in\mathbb{U}_\delta(0)\,\big\}\,.$$

Obviously, $\tau_\delta\leq\tau_0$. We will first show that for every $z\in\mathbb{U}_\delta(\rho^0)$,

$$\liminf_{\ell,m,q}\frac{1}{m}\ln E_z(\tau_\delta) \;\geq\; \inf\big\{\,V(r,t) : r\in U_\delta(\rho^0), t\in U_\delta(0)\,\big\} - \delta\,.$$

In order to ease the notation in the sequel, we set V^δ to be the infimum appearing in the above formula. Using Markov's inequality, for all $T\geq 0$

$$E_z(\tau_\delta) \;\geq\; TP_z(\tau_\delta\geq T)\,.$$

Thus, we set $T\geq 0$ and we bound the probability of the event $\{\,\tau_\delta<T\,\}$. Let us denote by T_0 the last time before τ_δ that the process is in $\mathbb{U}_\delta(\rho^0)$, i.e.,

$$T_0 \;:=\; \max\big\{\,n\leq\tau_\delta : Z_n\in\mathbb{U}_\delta(\rho^0)\,\big\}\,.$$

We will bound the probability of the event $\{\,\tau_\delta<T\,\}$ by studying the trajectory of the process $(Z_n)_{n\geq 0}$ between T_0 and τ_δ. The idea is the following, either the trajectory $(Z_n)_{T_0<n<\tau_\delta}$ spends a long time outside a neighborhood of ρ^0 and 0, which is very unlikely (Lemmas 4.1, 5.1 and Corollary 4.2), or it jumps in a few steps from one fixed point to another until reaching 0, in which case the lower bound of the Large Deviations Principle 3.2 will give us the desired estimate. We have, for $T\geq 0$, $z\in\mathbb{U}_\delta(\rho^0)$, and $k\leq T$,

$$P_z(\tau_\delta<T) \;=\; \sum_{0\leq t_0<t^*<T} P_z(T_0=t_0,\tau_\delta=t^*)$$
$$= \sum_{\substack{0\leq t_0<t^*<T\\ t^*-t_0\leq k}} P_z(T_0=t_0,\tau_\delta=t^*) + \sum_{\substack{0\leq t_0<t^*<T\\ t^*-t_0>k}} P_z(T_0=t_0,\tau_\delta=t^*)\,. \tag{6.6}$$

The first of the terms in the right–hand side of (6.6) can be bounded thanks to the Large Deviations Principle 3.3. Indeed, if $0 \leq t_0 < t^* < T$ and $t^* - t_0 < k$, we have

$$\begin{aligned} P_z\big(T_0 = t_0, \tau_\delta = t^*\big) \;&=\; \sum_{z' \in \mathbb{U}_\delta(\rho^0)} P_z\big(T_0 = t_0, \tau_\delta = t^*, Z_{t_0} = z'\big) \\ &\leq \sum_{z' \in \mathbb{U}_\delta(\rho^0)} P_{z'}\big(Z_{t^*-t_0} \in \mathbb{U}_\delta(0)\big) \;\leq\; m^{C(K)} \sup_{z' \in \mathbb{U}_\delta(\rho^0)} P_{z'}\big(Z_{t^*-t_0} \in \mathbb{U}_\delta(0)\big)\,, \end{aligned}$$

where $C(K)$ is a positive constant that depends on K but not on m. We have then,

$$\sum_{\substack{0 \leq t_0 < t^* < T \\ t^* - t_0 \leq k}} P_z\big(T_0 = t_0, \tau_\delta = t^*\big) \;\leq\; T m^{C(K)} \sum_{h=0}^{k} \sup_{z' \in \mathbb{U}_\delta(\rho^0)} P_{z'}\big(Z_h \in \mathbb{U}_\delta(0)\big).$$

Yet, thanks to the large deviations principle of Corollary 3.3,

$$\begin{aligned} &\limsup_{\ell,m,q} \frac{1}{m} \ln \sum_{h=0}^{k} \sup_{z' \in \mathbb{U}_\delta(\rho^0)} P_{z'}\big(Z_h \in \mathbb{U}_\delta(0)\big) \\ &\qquad\qquad \leq -\min_{0 \leq h \leq k} \inf\big\{\, V_h(x,y) : x \in U_\delta(\rho^0), y \in U_\delta(0) \,\big\} \;\leq\; -V^\delta\,. \end{aligned} \tag{6.7}$$

We deal next with the second term in the right–hand side of (6.6). Recall from Section 2.10 that U_δ represents the union of the δ–neighborhoods of all the fixed points and W_δ represents this same union where the neighborhoods of ρ^0 and 0 have been left out. We define the random time T_1^* by

$$T_1^* \;:=\; \min\big\{\, n \geq T_0 : Z_n \in \mathbb{W}_\delta \,\big\}\,.$$

We break the second term of the right–hand side of (6.6) as follows,

$$\begin{aligned} \sum_{\substack{0 \leq t_0 < t^* < T \\ t^* - t_0 > k}} P_z\big(T_0 = t_0, \tau_\delta = t^*\big) \;&=\; \sum_{\substack{0 \leq t_0 < t^* < T \\ t^* - t_0 > e^{\varepsilon m}}} P_z\big(T_0 = t_0, \tau_\delta = t^*\big) \\ &+ \sum_{\substack{0 \leq t_0 < t^* < T \\ k < t^* - t_0 < e^{\varepsilon m}}} P_z\big(T_0 = t_0, \tau_\delta = t^*, Z_t \notin \mathbb{U}_\delta, t_0 < t < t^*\big) \\ &+ \sum_{\substack{0 \leq t_0 < t_1^* < t^* < T \\ k < t^* - t_0 < e^{\varepsilon m}}} P_z\big(T_0 = t_0, T_1^* = t_1^*, \tau_\delta = t^*\big)\,. \end{aligned} \tag{6.8}$$

The first of the sums in the right–hand side of (6.8) can be bounded thanks to the inequality (6.5). Indeed, if $t^* - t_0 > e^{\varepsilon m}$, then

$$\begin{aligned} P_z\big(T_0 = t_0, \tau_\delta = t^*\big) \;&=\; \sum_{z' \in \mathbb{U}_\delta(\rho^0)} P_z\big(T_0 = t_0, Z_{t_0} = z', \tau_\delta = t^*\big) \\ &\leq \sum_{z' \in \mathbb{U}_\delta(\rho^0)} P_{z'}\big(Z_t \notin \mathbb{U}_\delta(\rho^0) \cup \mathbb{U}_\delta(0), 0 \leq t \leq t^* - t_0\big) \\ &\leq m^{C(K)} \sup_{z' \in \mathbb{U}_\delta(\rho^0)} P_{z'}\big(Z_t \notin \mathbb{U}_\delta(\rho^0) \cup \mathbb{U}_\delta(0), 0 \leq t \leq t^* - t_0\big) \\ &\leq m^{C(K)} e^{-\varepsilon m/2}\,. \end{aligned} \tag{6.9}$$

The second of the sums in the right–hand side of (6.8) can be bounded thanks to Corollary 4.2. Let h and c be as in Corollary 4.2. Then, we have, for $0 \leq t_0 < t^* < T$ and

$t^* - t_0 > k$,

$$\begin{aligned} P_z\big(T_0 = t_0, \tau_\delta = t^*, Z_t \notin \mathbb{U}_\delta, t_0 < t < t^*\big) \\ = \sum_{z' \in \mathbb{D} \setminus \mathbb{U}_\delta} P_z\big(T_0 = t_0, Z_{t_0+1} = z', \tau_\delta = t^*, Z_t \notin \mathbb{U}_\delta, t_0 < t < t^*\big) \\ \leq \sum_{z' \in \mathbb{D} \setminus \mathbb{U}_\delta} P_{z'}\big(Z_t \in \mathbb{D} \setminus \mathbb{U}_\delta, 0 \leq t < t^* - t_0\big) \\ \leq\; m^{C'(K)} \exp\left(-mc \Big\lfloor \frac{t^* - t_0}{h} \Big\rfloor \right), \end{aligned} \tag{6.10}$$

where $C'(K)$ is a positive constant that depends on K but not on m. Thus,

$$\sum_{\substack{0 \leq t_0 < t^* < T \\ t^* - t_0 > k}} P_z\begin{pmatrix} T_0 = t_0, \tau_\delta = t^*, \\ Z_t \notin \mathbb{U}_\delta, t_0 < t < t^* \end{pmatrix} \;\leq\; T^2 m^{C'(K)} \exp\left(-mc \Big\lfloor \frac{k}{h} \Big\rfloor \right).$$

In order to bound the last sum in (6.8), we introduce, for $b \in I(A)$, the random time

$$T_1^b \;:=\; \sup\big\{\, n \leq t^* : Z_n \in \mathbb{U}_\delta(\rho^b) \,\big\}\,.$$

We decompose the last term in (6.8) as follows,

$$\sum_{\substack{0 \leq t_0 < t_1^* < t^* < T \\ k < t^* - t_0 < e^{\varepsilon m}}} P_z\big(T_0 = t_0, T_1^* = t_1^*, \tau_\delta = t^*\big) \;=$$

$$\sum_{\substack{b \in I(A) \\ b \neq 0, K+1}} \; \sum_{\substack{0 \leq t_0 < t_1^* < t_1 < t^* < T \\ k < t^* - t_0 < e^{\varepsilon m}}} P_z\begin{pmatrix} T_0 = t_0, T_1^* = t_1^*, \\ Z_{t_1^*} \in \mathbb{U}_\delta(\rho^b), T_1^b = t_1, \tau_\delta = t^* \end{pmatrix}.$$

For a given $b \in I(A) \setminus \{\, 0, K+1 \,\}$, we decompose further the above sum, by considering the three following cases:

• If $t_1^* - t_0 > k$, then, for some positive constant $C(K)$ that depends on K only, the above sum can be bounded by

$$T e^{3\varepsilon m} m^{C(K)} \exp\left(-m \Big\lfloor \frac{k}{h} \Big\rfloor \right).$$

• If $t_1^* - t_0 < k$ and $t^* - t_1 < k$, then the sum can be bounded thanks to the large deviations principle in Corollary 3.3, which gives the following bound:

$$\begin{aligned} T e^{\varepsilon m} \exp\bigg(-m\Big(\inf\big\{\, V(x,y) : x \in U_\delta(\rho^0), y \in U_\delta(\rho^b) \,\big\} - \varepsilon \Big) \bigg) \times \\ \exp\bigg(-m\Big(\inf\big\{\, V(x,y) : x \in U_\delta(\rho^b), y \in U_\delta(0) \,\big\} - \varepsilon \Big) \bigg) < T e^{-m(V^\delta - 3\varepsilon)}\,. \end{aligned}$$

• If $t_1^* - t_0 < k$ and $t^* - t_1 > k$, then we define the set W_δ^{-b} by

$$W_\delta^{-b} \;:=\; \bigcup_{\substack{b' \in I(A) \\ b' \neq 0, b, K+1}} U_\delta(\rho^b)\,,$$

and we define the hitting time of $\mathbb{W}_\delta^{-b}$ after the time T_1 by

$$T_2^* \;:=\; \inf\big\{\, n \geq T_1 : Z_n \in \mathbb{W}_\delta^{-b} \,\big\}\,.$$

Then, the sum can be bounded by

$$\sum_{\substack{0\leq t_0<t_1^*<t_1<t \\ t^*-t_1>k,\, t^*-t_0<e^{\varepsilon m}}} P_z\begin{pmatrix} T_0=t_0, T_1^*=t_1^*, Z_{t_1^*}\in\mathbb{U}_\delta(\rho^b), \\ T_1=t_1, \tau_\delta=t^*, Z_t\notin\mathbb{U}_\delta, t_1<t<t_1^*\end{pmatrix}$$

$$+\sum_{\substack{0\leq t_0<t_1^*<t_1<t_2^*<t \\ t_1^*-t_0<k,\, t^*-t_0<e^{\varepsilon m}}} P_z\begin{pmatrix} T_0=t_0, T_1^*=t_1^*, Z_{t_1^*}\in\mathbb{U}_\delta(\rho^b), \\ T_1=t_1, T_2^*=t_2, \tau_\delta=t^*\end{pmatrix}$$

The first of the sums is again bounded by

$$Te^{3\varepsilon m}m^{C(K)}\exp\left(-mc\left\lfloor\frac{k}{h}\right\rfloor\right).$$

In order to bound the second sum, we can break it again in three different cases, and iterate this same procedure until we exhaust the fixed points in the set $I(A)$. We will then get $3|I(A)|$ summands, each of them being bounded by

$$\max\left\{Te^{M(K)\varepsilon m}m^{C(K)}\exp\left(-mc\left\lfloor\frac{k}{h}\right\rfloor\right), Te^{-m(V^\delta-M(K)\varepsilon)}\right\}, \tag{6.11}$$

where $M(K)$ is a natural number depending on K only. We choose k large enough so that

$$c\left\lfloor\frac{k}{h}\right\rfloor > \inf\left\{V(x,y): x\in U_\delta(\rho^0), y\in U_\delta(0)\right\}.$$

We set

$$T := \exp\left(m\Big(\inf\left\{V(x,y): x\in U_\delta(\rho^0), y\in U_\delta(0)\right\}-\delta\Big)\right).$$

Then, taking ε small enough so that $M(K)\varepsilon<\delta$, and combining the estimates (6.7), (6.9), (6.10), and (6.11), we conclude that, asymptotically,

$$P_z(\tau_\delta\leq T) \leq e^{-m(V_\delta-M(K)\varepsilon)}. \tag{6.12}$$

We deduce from here that, for every $z\in\mathbb{U}_\delta(\rho^0)$,

$$\liminf_{\ell,m,q}\frac{1}{m}\ln E_z(\tau_\delta)\geq\inf\left\{V(x,y): x\in U_\delta(\rho^0), y\in U_\delta(0)\right\}-\delta.$$

Now let $z\in\mathbb{D}\setminus\{0\}$ and note that, from Corollary 5.7, we can deduce that there exists $C>0$ such that

$$P_z\big(Z_{\lfloor C\ln m\rfloor}\in\mathbb{U}_\delta(\rho^0), Z_t\neq 0, 0\leq t\leq\lfloor C\ln m\rfloor\big) \geq e^{-\delta m}.$$

Therefore, for every $T\geq\lfloor C\ln m\rfloor$, we have

$$P_z(\tau_0>T) \geq \sum_{z'\in\mathbb{U}_\delta(\rho^0)} P_z\big(\tau_0>T, Z_{\lfloor C\ln m\rfloor}=z'\big) \geq$$
$$\sum_{z'\in\mathbb{U}_\delta(\rho^0)} P_z\big(Z_{\lfloor C\ln m\rfloor}=z', Z_t\neq 0, 0\leq t\leq\lfloor C\ln m\rfloor\big)P_{z'}\big(\tau_0>T-\lfloor C\ln m\rfloor\big).$$

Thus, for any $z\in\mathbb{D}\setminus\{0\}$

$$E_z(\tau_0) = \sum_{T\geq 0}P_z(\tau_0>T) \geq \sum_{T\geq\lfloor C\ln m\rfloor}P_z(\tau_0>T)$$
$$\geq \sum_{z'\in\mathbb{U}_\delta(\rho^0)} P_z\big(Z_{\lfloor C\ln m\rfloor}=z', Z_t\neq 0, 0\leq t\leq\lfloor C\ln m\rfloor\big)$$
$$\times\sum_{T\geq\lfloor C\ln m\rfloor} P_{z'}\big(\tau_0>T-\lfloor C\ln m\rfloor\big) \geq e^{-\varepsilon m}\inf_{z'\in\mathbb{U}_\delta(\rho^0)}E_{z'}(\tau_0).$$

Yet, $\tau_\delta < \tau_0$ by definition. Thus, for any $z \in \mathbb{D} \setminus \{ 0 \}$,

$$\liminf_{\ell,m,q} \frac{1}{m} \ln E_z(\tau_0) \;\geq\; \inf \big\{ V(x,y) : x \in U_\delta(\rho^0), y \in U_\delta(0) \big\} - 2\delta \,.$$

We let δ go to zero and we get the desired result. □

In fact, in view of (6.12), we can further establish the following concentration result for the time τ_0.

Corollary 6.2. *For every $\delta > 0$, there exists $\varepsilon > 0$ such that asymptotically,*

$$\forall\, z \in \mathbb{D} \setminus \{ 0 \} \,, \qquad P_z\big(\tau_0 \geq e^{m(V(\rho^0,0)-\delta)}\big) \;>\; 1 - e^{-\varepsilon m} \,.$$

7 Concentration near ρ^0

We assume throughout this whole section that $A(0)e^{-a} > 1$. Our purpose is to study the behavior of the process $(Z_n)_{n\geq 0}$ inside the set $\mathbb{D}$, in order to show that it spends most of its time close to the fixed point ρ^0. We introduce the following stopping times: set $T_0 := 0$ and

$$\begin{array}{ll} T_1^* := \inf\big\{n \geq T_0 : Z_n \in \mathbb{U}_\delta^0\big\} & T_1 := \inf\big\{n \geq T_1^* : Z_n \notin \mathbb{U}_{2\delta}^0\big\} \\ \quad\vdots & \quad\vdots \\ T_i^* := \inf\big\{n \geq T_{i-1} : Z_n \in \mathbb{U}_\delta^0\big\} & T_i := \inf\big\{n \geq T_i^* : Z_n \notin \mathbb{U}_{2\delta}^0\big\} \\ \quad\vdots & \quad\vdots \end{array}$$

Set also

$$\begin{aligned} \tau_0 \;&:=\; \inf\big\{ n \geq 0 : Z_n = 0 \big\} \,, \\ \iota(n) \;&:=\; \max\big\{ i \leq n : T_{i-1} < n \big\} \,. \end{aligned}$$

Our purpose is to show the following result.

Theorem 7.1. *For any $\varepsilon > 0$, we have*

$$E\bigg(\sum_{i=0}^{\iota(\tau_0)} (T_i^* \wedge \tau_0 - T_{i-1}) \bigg) \;\leq\; e^{m(V-\varepsilon)} \,.$$

Before jumping to the proof of the theorem, let us give a couple of auxiliary lemmas.

Lemma 7.2. *Let $\delta, \varepsilon > 0$. There exists $C = C(\delta,\varepsilon) > 0$ such that, asymptotically, for every $z \in \mathbb{D} \setminus \{ 0 \}$,*

$$\forall\, n \geq 0 \qquad P_z\big(T_1^* \wedge \tau_0 \geq n\lfloor C \ln m \rfloor\big) \;\leq\; (1 - e^{-\varepsilon m})^n \,.$$

The proof is similar to that of Corollary 4.2, using the estimate of Corollary 5.7.

Let τ denote the exit time of the process $(Z_n)_{n\geq 0}$ from the set $\mathbb{U}_{2\delta}^0$, i.e.,

$$\tau \;:=\; \inf\big\{ n \geq 0 : Z_n \notin \mathbb{U}_{2\delta}^0 \big\} \,.$$

We have the following bound on τ.

Lemma 7.3. *There exist $\gamma, \gamma' > 0$ such that, asymptotically, for all $z \in \mathbb{U}_\delta^0$,*

$$P_z\big(\tau \leq e^{m\gamma}\big) \;<\; e^{-\gamma' m} \,.$$

Proof. Define S to be the last time before τ that the process is in $\mathbb{U}_\delta^0$, i.e.,

$$S := \sup\big\{ 0 \le n \le \tau : Z_n \in \mathbb{U}_\delta^0 \big\} .$$

For any $n \ge 1$,

$$P_z\big(\tau \le n\big) = \sum_{0\le s<t\le n} P_z\big(S=s,\tau=t\big) .$$

Let $h = h \ge 2$ and $c = c > 0$ be as in Corollary 4.2. For a given value of s, we split the sum over t in two parts:

$$\sum_{t:s<t\le n} P_z\big(S=s,\tau=t\big) = \sum_{t:t>s+h+1} \cdots \quad + \sum_{t:s<t\le s+h+1} \cdots .$$

We study next the first sum, when $t > s + h + 1$. We condition on the state of the process at time $s + 1$. By the Markov property,

$$\begin{aligned}\sum_{t:t>s+h+1} \cdots = & \sum_{\substack{t:t>s+h+1\\ z'\in\mathbb{U}_{2\delta}^0\setminus\mathbb{U}_\delta^0}} P_z\big(S=s,Z_{s+1}=z',\tau=t\big)\\ & = \sum_{\substack{t:t>s+h+1\\ z'\in\mathbb{U}_{2\delta}^0\setminus\mathbb{U}_\delta^0}} P_z\begin{pmatrix} S=s,Z_{s+1}=z',\tau=t,\\ Z_{s+1},\dots,Z_{t-1}\in\mathbb{U}_{2\delta}^0\setminus\mathbb{U}_\delta^0\end{pmatrix}\\ & \le \sum_{\substack{t:t>s+h+1\\ z'\in\mathbb{U}_{2\delta}^0\setminus\mathbb{U}_\delta^0}} P_z\big(Z_{s+1}=z'\big)P_{z'}\big(Z_1,\dots,Z_{t-s-2}\in\mathbb{U}_{2\delta}^0\setminus\mathbb{U}_\delta^0, Z_{t-s-1}\notin\mathbb{U}_{2\delta}^0\big) .\end{aligned}$$

Since the set $\mathbb{U}_{2\delta}^0 \setminus \mathbb{U}_\delta^0$ contains none of the fixed points, and since $t > s + h + 1$, by Corollary 4.2, this last probability is smaller than $\exp(-mc\lfloor(t-s-2/h)\rfloor)$. Therefore,

$$\sum_{t:t>s+h+1} \cdots \le \sum_{t\ge h} e^{-mc\lfloor \frac{t}{h}\rfloor} = \frac{he^{-mc}}{1-e^{-mc}} .$$

We bound next the second sum. Conditioning on the state at time s:

$$\begin{aligned}\sum_{t:s<t\le s+h+1} \cdots = & \sum_{\substack{t:s<t\le s+h+1\\ z'\in\mathbb{U}_\delta^0}} P_z\big(S=s,Z_s=z',\tau=t\big)\\ \le & \sum_{\substack{t:s<t\le s+h+1\\ z'\in\mathbb{U}_\delta^0}} P_{z'}\big(Z_t\notin\mathbb{U}_{2\delta}^0\big)P_z\big(Z_s=z'\big) = \sum_{\substack{t:1\le t\le h+1\\ z'\in\mathbb{U}_\delta^0}} P_{z'}\big(Z_t\notin\mathbb{U}_{2\delta}^0\big)P_z\big(Z_s=z'\big) .\end{aligned}$$

Using the large deviation principle of Corollary 3.3, since h is fixed, for any $t \le h + 1$

$$\limsup_{\ell,m,q} \sup_{z'\in\mathbb{U}_\delta^0} P_{z'}\big(Z_t\notin\mathbb{U}_{2\delta}^0\big) \le -\inf\big\{ V(x,y) : x\in U_\delta^0, y\notin U_{2\delta}^0 \big\} .$$

Recall that δ has been chosen small enough so that $G(\overline{U_\delta^0}) \subset U_\delta^0$. Thus, the above infimum is strictly positive. We deduce that there exists $c' > 0$ (depending on δ) such that

$$\sum_{t:1\le t\le h} P_{z'}\big(Z_t\notin\mathbb{U}_{2\delta}^0\big) \le e^{-c'm} ,$$

the bound being uniform over $z' \in \mathbb{U}_\delta^0$. Finally, we obtain, for any $n \ge 1$,

$$P_z(\tau\le n) \le \sum_{1\le s\le n} \frac{he^{-mc}}{1-e^{-mc}} + \sum_{1\le s\le n} e^{-c'm} \le ne^{-c''m} ,$$

for some $c'' > 0$. Picking γ so that $\gamma' + \gamma - c'' < 0$, we have

$$P_z(\tau \leq e^{\gamma m}) \;\leq\; e^{-\gamma' m}\,,$$

as wanted. □

Proof of Theorem 7.1. Note that the argument in the expectation is bounded by τ_0, so for any $i^* \in \mathbb{N}$, we can break the above expectation as follows:

$$E\bigg(\sum_{i=1}^{\iota(\tau_0)} (T_i^* \wedge \tau_0 - T_{i-1})\bigg) \;\leq\; \sum_{i=1}^{i^*} E\big(1_{i \leq \iota(\tau_0)}(T_i^* - T_{i-1})\big) + E(\tau_0 1_{\iota(\tau_0) > i^*})\,.$$

If $1 \leq i \leq \iota(\tau_0)$, then $T_{i-1} \leq \tau_0$ and $Z_{T_{i-1}} \neq 0$, so that, using the Markov property,

$$E\big(1_{i \leq \iota(\tau_0)}(T_i^* - T_{i-1})\big) \;\leq\; \sup_{z \in \mathbb{D}\setminus\{0\}} E_z\big(T_1^* \wedge \tau_0\big)\,.$$

Thanks to the bound in Lemma 7.2, asymptotically, for any $z \in \mathbb{D} \setminus \{\,0\,\}$

$$\begin{aligned} E_z\big(T_1^* \wedge \tau_0\big) \;&=\; \sum_{k \geq 1} P_z\big(T_1^* \wedge \tau_0 \geq k\big) \\ &\leq \sum_{n \geq 0} \sum_{k = n\lfloor C \ln m \rfloor + 1}^{(n+1)\lfloor C \ln m \rfloor} P_z\big(T_1^* \wedge \tau_0 \geq n \lfloor C \ln m \rfloor\big) \\ &\leq\; \lfloor C \ln m \rfloor \sum_{n \geq 0} (1 - e^{-\varepsilon m})^n \;\leq\; e^{2\varepsilon m}\,. \end{aligned}$$

We conclude that, for any $i^* \in \mathbb{N}$ and $\varepsilon > 0$,

$$E\bigg(\sum_{i=0}^{\iota(\tau_0)} (T_i^* \wedge \tau_0 - T_{i-1})\bigg) \;\leq\; i^* e^{\varepsilon m} + E\big(\tau_0 1_{\iota(\tau_0 > i^*)}\big)\,.$$

Let $\eta > 0$ and define $t_m^\eta = e^{m(V+\eta)}$. Then,

$$\begin{aligned} E\big(\tau_0 1_{\iota(\tau_0) > i^*}\big) \;&=\; E\big(\tau_0 1_{\iota(\tau_0) > i^*} 1_{\tau_0 > t_m^\eta}\big) + E\big(\tau_0 1_{\iota(\tau_0) > i^*} 1_{\tau_0 \leq t_m^\eta}\big) \\ &\leq\; E\big(\tau_0 1_{\tau_0 > t_m^\eta}\big) + t_m^\eta P\big(\iota(t_m^\eta) > i^*\big) \end{aligned} \tag{7.1}$$

Let us begin by bounding the first term on the right–hand side of this inequality. We have, for every $n \in \mathbb{N}$ and $z \in \mathbb{D} \setminus \{\,0\,\}$,

$$\begin{aligned} E_z\big(\tau_0 1_{\tau_0 > n}\big) \;&=\; \sum_{k \geq 0} P_z\big(\tau_0 1_{\tau_0 > n} > k\big) \\ &= \sum_{k \geq 0} P_z\big(\tau_0 > k \vee n\big) \;\leq\; n P_z\big(\tau_0 > n\big) + \sum_{k \geq n} P_z\big(\tau_0 > k\big)\,. \end{aligned}$$

From inequality (6.4), for every $\gamma > 0$, there exists $C > 0$ such that

$$\forall\, h \geq 1 \quad \forall\, z \in \mathbb{D} \setminus \{\,0\,\} \qquad P_z\big(\tau_0 > h \lfloor C \ln m \rfloor\big) \;\leq\; \Big(1 - e^{-m(V+\gamma)}\Big)^h\,.$$

Using this inequality with $\gamma = \eta/2$, and setting $n = h\lfloor C \ln m \rfloor$, we get

$$\begin{aligned} E_z\big(\tau_0 1_{\tau_0 > n}\big) &\leq h \lfloor C \ln m \rfloor \Big(1 - e^{-m(V+\eta/2)}\Big)^h + \sum_{i \geq h} \lfloor C \ln m \rfloor P_z\big(\tau_0 > i \lfloor C \ln m \rfloor\big) \\ &\leq\; h \lfloor C \ln m \rfloor \Big(1 - e^{-m(V+\eta/2)}\Big)^h + \lfloor C \ln m \rfloor \Big(1 - e^{-m(V+\eta/2)}\Big)^h e^{m(V+\eta/2)}\,. \end{aligned}$$

Yet, if $h = \lfloor t_m^\eta / \lfloor C \ln m \rfloor \rfloor$, and since for m large $\lfloor C \ln m \rfloor \leq e^{\eta m/2}$, we have, asymptotically,

$$E_z\big(\tau_0 1_{\tau_0 > t_m^\eta}\big) \leq 2e^{m(V+\eta)}\Big(1 - e^{-m(V+\eta/2)}\Big)^h \leq 2e^{m(V+\eta)} e^{-h\exp(-mV - m\eta/2)}.$$

And since

$$he^{-m(V+\eta/2)} \;\geq\; \left(\frac{e^{m(V+\eta)}}{\lfloor C \ln m \rfloor} - 1\right) e^{-m(V+\eta/2)} \;\geq\; \frac{e^{m\eta/2}}{\lfloor C \ln m \rfloor} - 1\,,$$

the above expectation goes to 0 when m goes to infinity. We deal now with the second term in (7.1). We set $i^* = 2e^{m(V+\eta-\gamma)}$. Then, combining Lemma 7.3 and Lemma A.1, there exists $C > 0$ such that

$$t_m^\eta P\big(\iota(t_m^\eta) > i^*\big) \;=\; t_m^\eta P\big(\iota(\frac{i^*}{2} e^{\gamma m}) > i^*\big) \;<\; t_m^\eta e^{-(i^*-1)C}\,,$$

which goes to 0 when m goes to infinity. We conclude, by choosing $\eta, \gamma, \varepsilon$ such that $\eta - \gamma + \varepsilon < -\varepsilon$, that

$$E\bigg(\sum_{i=0}^{\iota(\tau_0)} (T_i^* \wedge \tau_0 - T_{i-1})\bigg) \;\leq\; e^{m(V-\varepsilon)}\,. \qquad \square$$

8 The neutral phase

The aim of this section is to study the process $(O_n)_{n\geq 0}$ when none of the classes $0, \dots, K$ are present in the population. Nevertheless, instead of using the occupancy process $(O_n)_{n\geq 0}$ for our study, we will use a related process, namely the distance process $(D_n)_{n\geq 0}$. The distance process is a Markov chain on $\{\,0, \dots, \ell\,\}^m$; an element $d \in \{\,0, \dots, \ell\,\}^m$, is a vector that represents the distances to the master sequence of the m individuals present in the population. The transition matrix p_H of the distance process is given by

$$\forall d, e \in \{\,0, \dots, \ell\,\}^m \qquad p_H(d, e) \;:=\; \prod_{1\leq i\leq m} \left(\sum_{1\leq j\leq m} \frac{A(d(j))M(d(j), e(i))}{A(d(1)) + \cdots + A(d(m))}\right). \tag{8.1}$$

The distance process and the occupancy process are related by a standard lumping procedure (cf. Section 4 of [2]). Let $k \geq 0$ We are interested in measuring the hitting time τ_k^* of the set of populations containing the classes $0, \dots, k$. Let us define, with a slight abuse of notation,

$$\begin{aligned}
\mathcal{W}_k^* \;&:=\; \big\{\, d \in \{\,0, \dots, \ell\,\}^m : d(i) \leq k \text{ for some } 1 \leq i \leq m \,\big\}\,, \\
\mathcal{N}_k \;&:=\; \big\{\, d \in \{\,0, \dots, \ell\,\}^m : d(i) > k \text{ for all } 1 \leq i \leq m \,\big\}\,.
\end{aligned}$$

The hitting time τ_k^* is then defined by

$$\tau_k^* \;:=\; \inf\big\{\, n \geq 0 : D_n \in \mathcal{W}_k^* \,\big\}\,.$$

The dynamics of the process D_n, started form any point in the set $\mathcal{N}_K$, and until the time τ_K^*, is the same as if the fitness landscape were neutral. Since we are ultimately interested in the hitting time τ_k^* for $k \geq K$, we will assume throughout the rest of this section that the fitness function A is constant and equal to 1.

Neutral hypothesis. Throughout this section we assume that $A(k) = 1$ for all $k \geq 0$.

The distance process has been studied in detail in Section 8 of [2]. Next we summarize the results therein that are of pertinent to our case.

Let $(Y_n)_{n\geq 0}$ be the Markov chain with state space $\{\,0,\dots,\ell\,\}$ and transition matrix the mutation matrix M defined in (2.1). The Markov chain $(Y_n)_{n\geq 0}$ is monotone. Let us denote by $\mathcal{B}$ the binomial law $\mathcal{B}(\ell,1-1/\kappa)$, i.e.

$$\forall b\in\{\,1,\dots,\ell\,\}\qquad \mathcal{B}(b)\;:=\;\binom{\ell}{b}\Big(1-\frac{1}{\kappa}\Big)^{b}\Big(\frac{1}{\kappa}\Big)^{\ell-b}\,. \tag{8.2}$$

The Markov chain $(Y_n)_{n\geq 0}$ is reversible with respect to the binomial law $\mathcal{B}$. The binomial law $\mathcal{B}$ concentrates exponentially fast around its mean $\ell_\kappa\,:=\,\ell(1-1/\kappa)$. We have the following results.

Lemma 8.1. *For all $b<\ell/2$, we have*

$$\frac{1}{\kappa^\ell}\Big(\frac{\ell}{2b}\Big)^{b}\;\leq\;\mathcal{B}(b)\;\leq\;\frac{\ell^b}{\kappa^{\ell-b}}\,.$$

This is Lemma 8.1 in [2].

Proposition 8.2. *For all $n\geq\sqrt{\ell}$, we have*

$$P\big(Y_n\geq\ln\ell\,\big|\,Y_0=k+1\big)\;\geq\;1-\exp\Big(-\frac{1}{2}(\ln\ell)^2\Big)\,.$$

Proposition 8.2 in [2], states the same result except for the starting point $k+1$ replaced by 0. Our proposition follows by the monotonicity of the chain $(Y_n)_{n\geq 0}$.

Let us come back to the distance process $(D_n)_{n\geq 0}$, whose state space is $\{\,1,\dots,\ell\,\}^m$ an its transition matrix is given by (8.1). We consider the following partial order on $\{\,1,\dots,\ell\,\}^m$,

$$d\preceq e\quad\Leftrightarrow\quad d(i)\leq e(i)\text{ for all }i\in\{\,1,\dots,m\,\}\,.$$

Proposition 8.3. *Under the neutral hypothesis:*

- *The distance process is monotone with respect to the partial order $\preceq$.*
- *If the law of D_0 has positive correlations, the law of D_n does too.*

We refer the reader to Section 5.2 in [2] for the details concerning the monotonicity of the distance process. In particular, the above results are proved in Corollary 5.6 and Proposition 5.8 therein. We have the following result concerning the discovery time of the master sequence.

Lemma 8.4. *For any $d\in\mathcal{N}_0$,*

$$\lim_{\ell,m,q}\frac{1}{\ell}\ln E(\tau_0^*\,|\,D_0=d)\;=\;\ln\kappa\,.$$

This is Proposition 8.6 of [2]. Finally, we will need the following upper bound on the hitting time τ_k^*.

Lemma 8.5. *If ℓ is large enough, for any $b\in\{k+1,\dots,\ell\}$,*

$$P\big(\tau_k^*\leq n\,\big|\,D_0=(b)^m\big)\;\leq\;nm\frac{\mathcal{B}(0)+\cdots+\mathcal{B}(k)}{\mathcal{B}(b)}\;\leq\;nm(k+1)\frac{\mathcal{B}(k)}{\mathcal{B}(b)}\,.$$

The proof of the first inequality is very similar to the proof of Lemma 10.15 of [1]. The second inequality is a consequence of $\mathcal{B}(k)$ being the largest term in the numerator, when ℓ is large enough.

Since for any $k\geq 0$, the set $\mathcal{W}_0^*$ is contained in the set $\mathcal{W}_k^*$, the hitting time τ_0 must be larger than the hitting time τ_k. The following lemma is an immediate consequence of these observations.

Lemma 8.6. *Asymptotically, for any $k \geq 0$, $\varepsilon > 0$ and $d \in \mathcal{N}_k$,*

$$E(\tau_k^* \,|\, D_0 = d) \;\leq\; \kappa^{\ell(1+\varepsilon)}\,.$$

Our next purpose is to find a lower bound for τ_k^*.

Lemma 8.7. *For any $k \geq 0$, $\varepsilon > 0$ and $d \in \mathcal{N}_k$,*

$$\liminf_{\ell,m,q} \frac{1}{\ell} \ln P(\tau_k^* > \kappa^{\ell(1-\varepsilon)} \,|\, D_0 = d) \;=\; 0\,.$$

Proof. The distance process is monotone, so we may suppose without loss of generality that the process starts from $(k{+}1)^m$. In the sequel, we follow closely the proof of the lower bound in the Proposition 8.6 in [2]. We may couple the mutation events of the distance process (c.f. Section 5.1 in [2]) with an array $\big\{\, U_n^{i,j},\ 1 \leq i \leq m,\ 1 \leq j \leq \ell,\ n \geq 1 \,\big\}$ via the mapping

$$\mathcal{M}_H : \{\, 0, \dots, \ell \,\} \times [0,1]^\ell \quad\longrightarrow\quad \{\, 0, \dots, \ell \,\}$$

defined by

$$\forall\, b \in \{\, 0, \dots, \ell \,\} \quad \forall\, u_1, \dots, u_\ell \in [0,1]^\ell$$
$$\mathcal{M}_H(b, u_1, \dots, u_\ell) \;=\; b - \sum_{k=1}^{b} 1_{u_k < q/(\kappa-1)} + \sum_{k=b+1}^{\ell} 1_{u_k > 1-q}\,.$$

The mapping $\mathcal{M}_H$ is such that if $U_1, \dots, U_\ell$ are i.i.d. random variables with distribution $Unif([0,1])$, then

$$b, c \in \{\, 0, \dots, \ell \,\} \qquad P\big(\mathcal{M}_H(b, U_1, \dots, U_\ell) = c\big) \;=\; M(b,c)\,.$$

Here, the random variables $U_n^{i,1}, \dots, U_n^{i,\ell}$ are responsible for the individual i in the population mutation (or not) in the step $n-1 \to n$. We define the following "good event":

$$\mathcal{E} \;:=\; \Big\{ \forall n \leq \ell^{3/4} \quad \forall i \in \{\, 1, \dots, m \,\},\ U_n^{i,1} > \frac{q}{\kappa - 1}, \dots, U_n^{i,k} > \frac{q}{\kappa - 1} \Big\} \tag{8.3}$$

If the event $\mathcal{E}$ occurs, until time $\ell^{3/4}$ none of the mutation events will create a sequence belonging to $\mathcal{W}_k^*$. Proceeding as in [2], we have

$$\begin{aligned} P\big(\tau_k^* > \kappa^{\ell(1-\varepsilon)} \,\big|\, D_0 = (k+1)^m\big) \;\geq\; & P\big(\tau_k^* > \kappa^{\ell(1-\varepsilon)} \,\big|\, D_0 = (\ln \ell)^m\big) \\ & \times P\big(D_{\ell^{3/4}} \geq (\ln \ell)^m,\ \mathcal{E} \,\big|\, D_0 = (k+1)^m\big)\,. \end{aligned} \tag{8.4}$$

Using the FKG inequality, the fact that the distance process started from $(k+1)^m$ has positive correlations, and Proposition 8.2, we can bound the second probability on the right hand side of (8.4) as follows

$$\begin{aligned} P\big(D_{\ell^{3/4}} \geq (\ln \ell)^m,\ \mathcal{E} \,\big|\, D_0 = (k+1)^m\big) &\geq P\big(D_{\ell^{3/4}} \geq (\ln \ell)^m\big) P\big(\mathcal{E}\big) \\ &\geq P\big(Y_{\ell^{3/4}} \geq \ln \ell\big)^m P\big(\mathcal{E}\big) \geq \Big(1 - e^{-\frac{1}{2}(\ln \ell)^2}\Big)^m \Big(1 - \frac{q}{\kappa - 1}\Big)^{km\ell^{3/4}}\,. \end{aligned} \tag{8.5}$$

As for the first probability on the right hand side of (8.4), let $\varepsilon' > 0$ and condition on the population at time ℓ^2:

$$\begin{aligned} P\big(\tau_k^* > \kappa^{\ell(1-\varepsilon)} \,\big|\, D_0 = (\ln \ell)^m\big) \;\geq\; & P\big(\tau_k^* > \kappa^{\ell(1-\varepsilon)} \,\big|\, D_0 = (\ell_k(1-\varepsilon'))^m\big) \\ & \times P\big(\tau_k^* > \ell^2,\ D_{\ell^2} \geq (\ell_k(1-\varepsilon'))^m \,\big|\, D_0 = (\ln \ell)^m\big)\,, \end{aligned} \tag{8.6}$$

where $\ell_\kappa = \ell(1-1/\kappa)$ is the mean of the binomial law $\mathcal{B}$. For the second probability on the right hand side of (8.6), we have

$$P\big(\tau_k^* > \ell^2,\ D_{\ell^2} \geq (\ell_k(1-\varepsilon'))^m \,\big|\, D_0 = (\ln \ell)^m\big) \\ \geq\ P\big(D_{\ell^2} \geq (\ell_k(1-\varepsilon'))^m \,\big|\, D_0 = (\ln\ell)^m\big) - P\big(\tau_k^* \leq \ell^2,\ \big|\, D_0 = (\ln\ell)^m\big)\,. \quad (8.7)$$

As in [2], there exists some $c(\varepsilon') > 0$ such that the first probability on the right hand side of (8.7) is bounded below by

$$P\big(D_{\ell^2} \geq (\ell_k(1-\varepsilon'))^m \,\big|\, D_0 = (\ln\ell)^m\big)\ \geq\ \big(1-e^{-c(\varepsilon')\ell}\big)^m\,. \quad (8.8)$$

Using Lemma 8.5 along with Lemma 8.1, we have

$$P\big(\tau_k^* \leq \ell^2 \,\big|\, D_0 = (\ln\ell)^m\big) \leq (k+1)\ell^2 m\frac{\mathcal{B}(k)}{\mathcal{B}(\ln\ell)} \leq (k+1)\kappa^k\ell^{k+2}m\Big(\frac{2\ln\ell}{\ell}\Big)^{\ln\ell}\,. \quad (8.9)$$

As for the first probability on the right hand side of (8.6), we use Lemma 8.5 with $n=\kappa^{\ell(1-\varepsilon)}$ and $b=\ell_\kappa(1-\varepsilon)$, along with a standard large deviations estimate in order to conclude that, if ε' is chosen small enough, there exists $c(\varepsilon)>0$ such that, for ℓ large enough

$$P\big(\tau_k^* \leq \kappa^{\ell(1-\varepsilon)} \,\big|\, D_0 = (\ell_\kappa(1-\varepsilon'))^m\big) \\ \leq\ (k+1)\kappa^{\ell(1-\varepsilon)} m\frac{\mathcal{B}(k)}{\mathcal{B}(\ell_\kappa(1-\varepsilon'))}\ \leq\ (k+1)me^{-c(\varepsilon)\ell}\,. \quad (8.10)$$

Combining (8.5), (8.8), (8.9) and (8.10), we get

$$P\big(\tau_k^* > \kappa^{\ell(1-\varepsilon)} \,\big|\, D_0 = (k+1)^m\big)\ \geq\ \Big(1-e^{-\frac{1}{2}(\ln\ell)^2}\Big)^m\Big(1-\frac{q}{\kappa-1}\Big)^{km\ell^{3/4}} \\ \Big(1-(k+1)me^{-c(\varepsilon)\ell}\ \Big)\bigg(\big(1-e^{-c(\varepsilon')\ell}\big)^m - (k+1)\kappa^k\ell^{k+2}m\Big(\frac{2\ln\ell}{\ell}\Big)^{\ln\ell}\bigg)\,.$$

Finally, taking logarithms, dividing by ℓ and sending ℓ to infinity gives the desired result. □

9 Proof of Theorem 2.7

The aim of this section is to prove the Theorem 2.7. Recall that $V(\rho^0,0)$ represents the quasipotential from ρ^0 to 0 (cf. the formula (6.1)).

Definition 9.1. *Define the function $a\mapsto\psi(a)$ to be equal to $V(\rho^0,0)$ on $]0,\ln A(0)[$, and to be equal to 0 elsewhere.*

Note that this definition of the mapping ψ corresponds to the one given in Subsection 2.7; on one hand, when $a\in]0,\ln A(0)[$ we have $\rho(a)=\rho^0$ and (2.11) is just the quasipotential $V(\rho^0,0)$. On the other hand, when $a\in]\ln A(0),\infty[$ we have $\rho(a)=0$ so that we in (2.11) we have both $s^0=0$ and $s^l=0$, so that $\psi(a)=0$ follows from the identity $I_K(0,0)=0$.

We will first look at the subcritical case, i.e. the case $\alpha\psi(a)>\ln\kappa$, and we'll deal with the supercritical case afterwards.

9.1 The subcritical case

We suppose that $\alpha\psi(a)>\ln\kappa$, so that in particular, $A(0)e^{-a}>1$, and ρ^0 is well defined. Recall that the aim is to show that, for any continuous and bounded function

$f : \mathbb{R}^{K+1} \longrightarrow \mathbb{R}$

$$\lim_{\substack{\ell,m\to\infty,\, q\to 0\\ \ell q\to a,\, m/\ell\to\alpha}} \left| \int_{\mathcal{P}^m_{\ell+1}} f\Big(\frac{\pi_k(o)}{m}\Big) d\mu(o) - f(\rho^0) \right| \,=\, 0\,.$$

Let $f : \mathbb{R}^{K+1} \longrightarrow \mathbb{R}$ be a continuous, bounded function. By the Ergodic Theorem for Markov chains,

$$\left| \int_{\mathcal{P}^m_{\ell+1}} f\Big(\frac{\pi_k(o)}{m}\Big) d\mu(o) - f(\rho^0) \right| \,\leq\, \lim_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1} \left| f\Big(\frac{\pi_K(O_t)}{m}\Big) - f(\rho^0) \right|. \tag{9.1}$$

Let $\varepsilon > 0$. We will prove that this last quantity is smaller than ε, for m, ℓ large enough, q small enough, and $\ell q, m/\ell$ close enough to a, α. We break the state space $\mathcal{P}^m_{\ell+1}$ into two disjoint subsets, $\mathcal{W}^*_K$ (the populations containing at least an individual in one of the classes $0, \dots, K$) and $\mathcal{N}_K$ (the populations containing no individuals in any of the classes $0, \dots, K$). The process $(O_n)_{n\geq 0}$ will jump between these two sets. We define the following sequence of stopping times, we set $\tau_0 := 0$ and

$$\begin{aligned}
&\tau^*_1 := \inf\big\{ n \geq 0 : O_n \in \mathcal{W}^*_K \big\} && \tau_1 := \inf\big\{ n \geq \tau^*_1 : O_n \in \mathcal{N}_K \big\}\\
&\quad\vdots && \quad\vdots\\
&\tau^*_k := \inf\big\{ n \geq \tau_{k-1} : O_n \in \mathcal{W}^*_K \big\} && \tau_k := \inf\big\{ n \geq \tau^*_k : O_n \in \mathcal{N}_K \big\}\\
&\quad\vdots && \quad\vdots
\end{aligned}$$

Recall from Section 2.10 that U^0_δ represents the δ–neighborhood of ρ^0. The set $\mathcal{D}$ being compact, the function f is uniformly continuous on $\mathcal{D}$. We choose δ small enough so that for every $r \in U^0_{2\delta}$,

$$\big| f(r) - f(\rho^0) \big| \,<\, \varepsilon\,,$$

and so that the set U^0_δ satisfies $G(\overline{U^0_\delta}) \subset U^0_\delta$ (cf. Proposition 2.6). For each $k \geq 0$ we define the following sequence of stopping times, we set $T_{k,0} := \tau^*_k$ and

$$\begin{aligned}
&T^*_{k,1} := \inf\big\{ n \geq T_{k,0} : Z_n \in \mathbb{U}^0_\delta \big\} && T_{k,1} := \inf\big\{ n \geq T^*_{k,1} : Z_n \notin \mathbb{U}^0_{2\delta} \big\}\\
&\quad\vdots && \quad\vdots\\
&T^*_{k,i} := \inf\big\{ n \geq T_{k,i-1} : Z_n \in \mathbb{U}^0_\delta \big\} && T_{k,i} := \inf\big\{ n \geq T^*_{k,i} : Z_n \notin \mathbb{U}^0_{2\delta} \big\}\\
&\quad\vdots && \quad\vdots
\end{aligned}$$

We distinguish between three different situations: either O_n is in $\mathcal{N}_K$, or O_n is in $\mathcal{W}^*_K$ and $\pi_K(O_n)$ is inside $\mathbb{U}^0_{2\delta}$, or O_n is in $\mathcal{W}^*_K$ and $\pi_K(O_n)$ is outside $\mathbb{U}^0_{2\delta}$. We bound the sum in (9.1) by breaking it according to these three situations, which gives the following bound,

$$\begin{aligned}
\sum_{t=0}^{n-1} \left| f\Big(\frac{\pi_K(O_t)}{m}\Big) - f(\rho^0) \right| \,\leq\, & 2||f||_\infty \sum_{k\geq 1} \big(\tau^*_k \wedge n - \tau_{k-1} \wedge n \big)\\
&+ \varepsilon n + 2||f||_\infty \sum_{k\geq 1} \sum_{i\geq 1} \big(T^*_{k,i} \wedge \tau_k \wedge n - T_{k,i-1} \wedge \tau_k \wedge n \big)\,.
\end{aligned} \tag{9.2}$$

The next step is to bound the above sums. We start with the first one of them. We define, for $n \geq 1$, the random variable $\iota(n)$ by

$$\iota(n) \,:=\, \max\big\{\, k \geq 0 : \tau_{k-1} < n \,\big\}\,.$$

We can rewrite the sum with the help of this new random variable as

$$\sum_{k\geq 1}(\tau_k^* \wedge n - \tau_{k-1} \wedge n) \,=\, \sum_{k=1}^{\iota(n)}(\tau_k^* \wedge n - \tau_{k-1})\,. \tag{9.3}$$

Define by $\tau(\mathcal{N}_K)$ the hitting time of $\mathcal{N}_K$, i.e.,

$$\tau(\mathcal{N}_K) \,:=\, \inf\big\{\, n\geq 0 : O_n \in \mathcal{N}_K \,\big\}\,.$$

By Corollary 6.2, there exists a number $\gamma > 0$ such that,

$$\max_{o\in\mathcal{W}_K^*} P_o\big(\tau(\mathcal{N}_K) < e^{m(V-\varepsilon)}\big) \,\leq\, e^{-\gamma m}\,.$$

Thus, applying Lemma A.1 with $A = B = \mathcal{W}_K^*$, $N = e^{m(V-\varepsilon)}$, $p = e^{-\gamma m}$ and $\lambda = 1/2$, it follows that, for every $h \geq 2$,

$$P\Big(\iota\big(\frac{h}{2}e^{m(V-\varepsilon)}\big) \geq h\Big) \,\leq\, e^{-(h-1)c}\,,$$

where c is a positive constant, independent of h. The next step is to bound the quantity on the right hand side of (9.3). Since this quantity is obviously bounded by n, for any $i \geq 0$, we can decompose it according to whether $\iota(n)$ is greater or smaller than i and bound it as follows,

$$E\left(\sum_{k=1}^{\iota(n)}(\tau_k^* \wedge n) - \tau_{k-1}\right) \,\leq\, nP\big(\iota(n)\geq i\big) + \sum_{k=1}^{i} E\big(\tau_k^* - \tau_{k-1}\big)\,.$$

In view of the Lemma 8.6, asymptotically, for every $o \in \mathcal{N}_k$, we have $E_o(\tau_k^* - \tau_{k-1}) \leq \kappa^{\ell(1+\varepsilon)}$; we deduce that

$$E\left(\sum_{k=1}^{\iota(n)}(\tau_K^* \wedge n) - \tau_{k-1}\right) \,\leq\, nP\big(\iota(n)\geq i\big) + i\kappa^{\ell(1+\varepsilon)}\,.$$

Let us set

$$i_n \,:=\, \min\Big\{\, i : n \leq \frac{ie^{m(V-\varepsilon)}}{2} \,\Big\}\,.$$

On one hand, for $i = i_n$ we get

$$nP\big(\iota(n)\geq i_n\big) \,\leq\, nP\bigg(\iota\Big(\frac{i_n e^{m(V-\varepsilon)}}{2}\Big) \geq i_n\bigg) \,\leq\, \frac{i_n e^{m(V-\varepsilon)}}{2} e^{-(i_n-1)c}\,.$$

This quantity goes to 0 as n goes to ∞. On the other hand,

$$\frac{1}{n} i_n \kappa^{\ell(1+\varepsilon)} \,\leq\, \frac{2i_n\kappa^{\ell(1+\varepsilon)}}{(i_n-1)e^{m(V-\varepsilon)}}\,.$$

When n goes to ∞, this last quantity converges to

$$2\frac{\kappa^{\ell(1+\varepsilon)}}{e^{m(V-\varepsilon)}} \,=\, 2\exp\Big(-m\big(V-\varepsilon-\frac{\ell}{m}(1+\varepsilon)\ln\kappa\big)\Big)\,,$$

which, since we are looking at the subcritical case $\alpha\psi(a) > \ln\kappa$, for ε small enough, goes to 0 when ℓ, m go to ∞ and q goes to 0. We proceed next to bound the second of the sums in (9.2). For $n, k \geq 1$, we define the following random variables:

$$\iota^*(n) \,:=\, \max\big\{\, k\geq 0 : \tau_k^* \leq n \,\big\}\,,$$
$$\iota_k(n) \,:=\, \max\big\{\, i\geq 0 : T_{k,i-1} < n \,\big\}$$

We can rewrite the sum with the help of these new random variables as follows:

$$\sum_{k\geq 1}\sum_{i\geq 1}\big(T^*_{k,i}\wedge\tau_k\wedge n - T_{k,i-1}\wedge\tau_k\wedge n\big) \;=\; \sum_{k=1}^{\iota^*(n)-1}\left(\sum_{i=1}^{\iota_k(\tau_k)}\big(T^*_{k,i}\wedge\tau_k - T_{k,i-1}\big)\right)$$
$$+\sum_{i=1}^{\iota_{\iota^*(n)}(\tau_{\iota^*(n)})}\big(T^*_{\iota_{\iota^*(n)},i}\wedge\tau_{\iota^*_n}\wedge n - T_{\iota_{\iota^*(n)},i-1}\wedge n\big)\,.$$

Taking the expectation, the above sum can be bounded by

$$nP\big(\iota^*(n)\geq i_n\big)+\sum_{k=1}^{i_n}E\left(\sum_{i=1}^{\iota_k(\tau_k)}\big(T^*_{k,i}\wedge\tau_k - T_{k,i-1}\big)\right).$$

Noting that $\iota^*(n)\leq\iota(n)$, the first term can be shown to converge to 0 as n goes to ∞, as in the previous section. Let us deal with the expectation. We introduce the following stopping times: set $T_0=0$ and

$$\begin{aligned}
T^*_1 &:= \inf\big\{n\geq T_0 : Z_n\in\mathbb{U}^0_\delta\big\} & T_1 &:= \inf\big\{n\geq T^*_1 : Z_n\notin\mathbb{U}^0_{2\delta}\big\}\\
&\;\vdots & &\;\vdots\\
T^*_i &:= \inf\big\{n\geq T_{i-1} : Z_n\in\mathbb{U}^0_\delta\big\} & T_i &:= \inf\big\{n\geq T^*_i : Z_n\notin\mathbb{U}^0_{2\delta}\big\}\\
&\;\vdots & &\;\vdots
\end{aligned}$$

Set also

$$\begin{aligned}
\tau_0 &:= \inf\big\{\,n\geq 0 : Z_n=0\,\big\}\,,\\
\iota(n) &:= \max\big\{\,i\leq n : T_{i-1}<n\,\big\}\,.
\end{aligned}$$

Fix $k\in\{\,1,\dots,i_n\,\}$. By the Markov property,

$$E\left(\sum_{i=1}^{\iota_k(\tau_k)}\big(T^*_{k,i}\wedge\tau_k - T_{k,i-1}\big)\right)$$
$$=\sum_{z\in\mathbb{D}\setminus\{0\}}E\left(\sum_{i=1}^{\iota_k(\tau_k)}\big(T^*_{k,i}\wedge\tau_k - T_{k,i-1}\big)\,\middle|\,Z_{\tau^*_k}=z\right)P\big(Z_{\tau^*_k}=z\big)$$
$$\leq\sup_{z\in\mathbb{D}\setminus\{0\}}E_z\left(\sum_{i=1}^{\iota(\tau_0)}\big(T^*_i\wedge\tau_0 - T_{i-1}\big)\right).$$

Yet, by Theorem 7.1, the last expectation is bounded by $e^{m(V-\gamma)}$, for any $\gamma>0$. Therefore,

$$\frac{1}{n}\sum_{1\leq k\leq i_n}E\left(\sum_{i=1}^{\iota_k(\tau_k)}\big(T^*_{k,i}\wedge\tau_k - T_{k,i-1}\big)\right)\;\leq\;\frac{i_n}{n}e^{m(V-\gamma)}$$
$$\leq\;\frac{i_n-1}{n}e^{m(V-\gamma)}+\frac{e^{m(V-\gamma)}}{n}\;\leq\;\frac{2e^{m(V-\gamma)}}{e^{m(V-\varepsilon)}}+\frac{e^{m(V-\gamma)}}{n}\,.$$

The last term goes to 0 when n goes to ∞. And choosing $\varepsilon<\gamma$, the first one converges to 0 when m goes to ∞.

9.2 The supercritical case

We suppose that $\alpha\psi(a) < \ln\kappa$. Recall that the aim is to show that, for any continuous and bounded function $f : \mathbb{R}^{K+1} \longrightarrow \mathbb{R}$

$$\lim_{\substack{\ell,m\to\infty,\, q\to 0\\ \ell q\to a,\, m/\ell\to\alpha}} \left| \int_{\mathcal{P}^m_{\ell+1}} f\left(\frac{\pi_k(o)}{m}\right) d\mu(o) - f(0)\right| \,=\, 0\,.$$

Let $f : \mathbb{R}^{K+1} \longrightarrow \mathbb{R}$ be a continuous, bounded function. By the Ergodic Theorem for Markov chains,

$$\left| \int_{\mathcal{P}^m_{\ell+1}} f\left(\frac{\pi_k(o)}{m}\right) d\mu(o) - f(0)\right| \,\leq\, \lim_{n\to\infty} \frac{1}{n}\sum_{t=0}^{n-1} \left| f\left(\frac{\pi_K(O_t)}{m}\right) - f(0)\right| .$$

Proceeding as in the subcritical case, we obtain the following bound:

$$\sum_{t=0}^{n-1} \left| f\left(\frac{\pi_K(O_t)}{m}\right) - f(0)\right| \,\leq\, 2||f||_\infty \left(\sum_{k=1}^{\iota(n)-1} (\tau_k - \tau_k^*) + n - \tau^*_{\iota(n)}\right).$$

Denote by $\tau(\mathcal{W}^*_K)$ the hitting time of the set $\mathcal{W}^*_K$, i.e.,

$$\tau(\mathcal{W}^*_K) \,:=\, \big\{\, n\geq 0 : O_n \in \mathcal{W}^*_K \,\big\}\,.$$

In view of Lemma 8.7, for every $\varepsilon,\gamma > 0$, we have

$$\max_{o\in\mathcal{N}_K} P_o\big(\tau(\mathcal{W}^*_K) > \kappa^{\ell(1-\varepsilon)}\big) \,\geq\, e^{-\gamma m}\,.$$

Thus, using Lemma A.1 with $A = B = \mathcal{N}_k$, $\lambda = 1 - e^{-\gamma m}/2$ and $N = \kappa^{\ell(1-\varepsilon)}$, we conclude that, for all $h \geq 2$,

$$P\Big(\iota\big(h(1-e^{-\gamma m}/2)\kappa^{\ell(1-\varepsilon)}\big) \geq h\Big) \,<\, e^{-(h-1)c}\,, \tag{9.4}$$

where c is a positive constant which does not depend on h. The next step is to bound the quantity

$$E\left(\sum_{k=1}^{\iota(n)-1} \big(\tau_k - \tau_k^*\big) + n - \tau^*_{\iota(n)}\right).$$

Let $i \geq 1$. Since this quantity is obviously bounded by n, we can decompose it as according to whether $\iota(n)$ is greater or smaller than i and bound it as follows

$$E\left(\sum_{k=1}^{\iota(n)-1} \big(\tau_k - \tau_k^*\big) + n - \tau^*_{\iota(n)}\right) \,\leq\, nP\big(\iota(n)\geq i\big) + \sum_{k=1}^{i} E\big(\tau_k - \tau_k^*\big)\,.$$

Since for every $o \in \mathcal{W}^*_K$, by Theorem 6.1,

$$E\big(\tau_1 \,\big|\, O_0 = o\big) \,\leq\, \exp\big(m(V+\varepsilon)\big)\,,$$

we deduce that

$$E\left(\sum_{k=1}^{\iota(n)-1} \big(\tau_k - \tau_k^*\big) + n - \tau^*_{\iota(n)}\right) \,\leq\, nP\big(\kappa(n)\geq i\big) + i\exp\Big(m\big(\psi(a)+\varepsilon\big)\Big)\,.$$

Let us set

$$i_n \,:=\, \min\big\{\, i : n \leq i(1-e^{-\gamma m}/2)\kappa^{\ell(1-\varepsilon)}\,\big\}\,.$$

On one hand, for $i = i_n$, using (9.4), we get

$$nP\big(\iota(n) \geq i_n\big) \leq i_n(1 - e^{-\gamma m}/2)\kappa^{\ell(1-\varepsilon)} \times P\Big(\iota\big(i_n(1 - e^{-\gamma m}/2)\kappa^{\ell(1-\varepsilon)}\big) \geq i_n\Big) \leq i_n(1 - e^{-\gamma m}/2)\kappa^{\ell(1-\varepsilon)} e^{-(i_n-1)c}.$$

This quantity goes to 0 as n goes to infinity. On the other hand,

$$\frac{i_n}{n} e^{m(V+\varepsilon)} \leq \frac{i_n}{(i_n - 1)(1 - e^{-\gamma m}/2)\kappa^{\ell(1-\varepsilon)}} e^{m(V+\varepsilon)}.$$

When n goes to infinity, this last quantity converges to

$$\frac{e^{m(V+\varepsilon)}}{(1 - e^{-\gamma m}/2)\kappa^{\ell(1-\varepsilon)}} = \exp\Big(- \ell\big(\ln \kappa - \alpha\psi(a) - \varepsilon - \varepsilon \ln \kappa + (\ln(1 - e^{-\gamma m}/2))/\ell\big)\Big),$$

which, for ε small enough, goes to 0 with ℓ, m, q.

A Bounds on hitting times

Let E be a finite set and $(X_n)_{n\geq 0}$ a recurrent Markov chain on E. For a set $A \subset E$ we denote by τ_A the hitting time of A, i.e.,

$$\tau_A := \inf\{ n \geq 0 : X_n \in A \}.$$

Let $A \subset B \subset E$ and define the following sequence of stopping times, we set $T_0 = 0$ and

$$\begin{aligned} &T_1^* := \inf\{n \geq 0 : X_n \in A\} && T_1 := \inf\{n \geq T_1^* : X_n \notin B\} \\ &\quad\vdots && \quad\vdots \\ &T_k^* := \inf\{n \geq T_{k-1} : X_n \in A\} && T_k := \inf\{n \geq T_k^* : X_n \notin B\} \\ &\quad\vdots && \quad\vdots \end{aligned}$$

Define, for $n \geq 1$, the random variable $\iota(n)$ by

$$\iota(n) := \max\{ k \geq 0 : T_{k-1} < n \}.$$

Our objective is to give a bound on the random variable $\iota(n)$. Let us assume that there exist $N, p > 0$ such that

$$\max_{z\in A} P\big(\tau_{E\setminus B} \leq N \,\big|\, X_0 = z\big) < p.$$

Lemma A.1. *For any $h \geq 1$ and $\lambda > p$, there exists $c > 0$ (depending on λ but not on h), such that*

$$P\Big(\iota\big(h\lambda N\big) \geq h\Big) < e^{-(h-1)c}.$$

Proof. Let us assume that $h\lambda$ is an integer number (otherwise we may replace it by $\lfloor h\lambda \rfloor$). From the definition of $\iota(n)$, we see that

$$\iota\Big(h\lambda N\Big) \geq h \quad\Leftrightarrow\quad T_{h-1} < h\lambda N.$$

We define the random variables $(Y_i)_{i\geq 1}$ by setting

$$Y_i := T_i - T_i^*, \quad i \geq 1.$$

Then,

$$T_{h-1} \geq Y_1 + \cdots + Y_{h-1}.$$

In view of the assumption on $\tau_{E\setminus B}$, for every $i \geq 1$,

$$P\big(Y_i \leq N\big) < p.$$

We define the following sequence of Bernoulli random variables

$$\varepsilon_i := 1_{Y_i \leq N}, \quad i \geq 1.$$

Thus, if $T_{h-1} < h\lambda N$, at least $(h-1)\lambda$ of the random variables $Y_1, \dots, Y_{h-1}$ must satisfy $Y_i \leq N$. Whence,

$$P\big(T_{h-1} < h\lambda N\big) \leq P\big(\varepsilon_1 + \cdots + \varepsilon_{h-1} \geq (h-1)\lambda\big).$$

We use the exponential Chebyshev inequality in order to bound the last probability: for any $\beta > 0$ we have

$$P\big(\varepsilon_1 + \cdots + \varepsilon_{h-1} \geq (h-1)\lambda\big) \leq e^{-\beta\lambda} E\big(e^{\beta\varepsilon_1/(h-1)} \cdots e^{\beta\varepsilon_{h-1}/(h-1)}\big).$$

The random variables $\varepsilon_1, \dots, \varepsilon_{h-2}$ are measurable with respect to $\big(X_n, 0 \leq n \leq T^*_{h-1}\big)$. Thus, thanks to the strong Markov property,

$$\begin{aligned} E\big(\varepsilon^{\beta\varepsilon_1/(h-1)} \cdots e^{\beta\varepsilon_{h-1}/(h-1)}\big) \\ &= E\Big(E\big(e^{\beta\varepsilon_1/(h-1)} \cdots e^{\beta\varepsilon_{h-1}/(h-1)} \,\big|\, X_0, \dots, X_{T^*_{h-1}}\big)\Big) \\ &= E\Big(e^{\beta\varepsilon_1/(h-1)} \cdots e^{\beta\varepsilon_{h-2}/(h-1)} E\big(e^{\beta\varepsilon_{h-1}/(h-1)} \,\big|\, X_0, \dots, X_{\tau^*_{h-1}}\big)\Big). \end{aligned}$$

Yet, for all $x \in A$,

$$E\big(e^{\beta\varepsilon_1/(h-1)} \,\big|\, X_0 = z\big) \leq e^{\beta/(h-1)} p + 1 - p.$$

Iterating, this procedure, we obtain

$$E\big(e^{\beta\varepsilon_1/(h-1)} \cdots e^{\beta\varepsilon_{h-1}/(h-1)}\big) \leq \big(e^{\beta/(h-1)} p + 1 - p\big)^{h-1}.$$

We make the change of variables $\beta \to (h-1)\beta$ in order to obtain

$$P\big(\varepsilon_1 + \cdots + \varepsilon_{h-1} \geq (h-1)\lambda\big) \leq \exp\Big(-(h-1)\big(\beta\lambda - \ln(e^\beta p + 1 - p)\big)\Big).$$

Denote by $\Lambda^*(t)$ the Cramèr transform of the Bernoulli law with parameter p,

$$\Lambda^*(t) := \sup_{\beta \geq 0} \big(\beta t - \ln(e^\beta p + 1 - p)\big) = t \ln \frac{t}{p} + (1-t) \ln \frac{1-t}{1-p}.$$

Optimizing the previous inequality over β, we obtain

$$P\big(\varepsilon_1 + \cdots + \varepsilon_{h-1} \geq (h-1)\lambda\big) \leq \exp\big(-(h-1)\Lambda^*(\lambda)\big),$$

where $\Lambda^*(\lambda) > 0$ is independent of h. It follows that

$$P\big(\iota(h\lambda N) \geq h\big) \leq e^{-(h-1)\Lambda^*(\lambda)},$$

as wanted. □

References

[1] Raphaël Cerf, *Critical population and error threshold on the sharp peak landscape for a Moran model*, Mem. Amer. Math. Soc. **233** (2015), no. 1096, vi+87. MR3307514

[2] Raphaël Cerf, *Critical population and error threshold on the sharp peak landscape for the Wright-Fisher model*, Ann. Appl. Probab. **25** (2015), no. 4, 1936–1992. MR3348998

[3] Raphaël Cerf and Joseba Dalmau, *Quasispecies on class-dependent fitness landscapes*, Bull. Math. Biol. **78** (2016), no. 6, 1238–1258. MR3523138

[4] Joseba Dalmau, *The distribution of the quasispecies for the Wright–Fisher model on the sharp peak landscape*, Stochastic Processes and their Applications **125** (2015), no. 1, 272 – 293. MR3274700

[5] Joseba Dalmau, *Asymptotic behavior of Eigen's quasispecies model*, Bull. Math. Biol. **80** (2018), no. 7, 1689–1712. MR3814758

[6] E. Domingo and P. Schuster, *Quasispecies: From theory to experimental systems*, Current Topics in Microbiology and Immunology, Springer International Publishing, 2016.

[7] Manfred Eigen, *Self-organization of matter and the evolution of biological macromolecules*, Naturwissenschaften **58** (1971), no. 10, 465–523.

[8] Manfred Eigen and Schuster Peter, *The hypercycle. A principle of natural self–organization. Part A: Emergence of the hypercycle*, Naturwissenschaften **64** (1977), no. 11, 541 – 565.

[9] Mark I. Freidlin and Alexander D. Wentzell, *Random perturbations of dynamical systems*, third ed., Springer, Heidelberg, 2012. MR2953753

[10] Selma Gago, Santiago F. Elena, Ricardo Flores, and Rafael Sanjuán, *Extremely high mutation rate of a hammerhead viroid*, Science **323** (2009), no. 5919, 1308–1308.

[11] S. R. S. Varadhan, *Large deviations*, Ann. Probab. **36** (2008), no. 2, 397–419. MR2393987

Acknowledgments. The autor is grateful to Raphaël Cerf for his helpful comments as well as to an anonymous referee.