



SURVIVAL OF THE FLATTEST IN THE QUASISPECIES MODEL

MAXIME BERGER,* *Université de Bourgogne*
RAPHAËL CERF,** *Université Paris-Saclay*

Abstract

Viruses present an amazing genetic variability. An ensemble of infecting viruses, also called a viral quasispecies, is a cloud of mutants centered around a specific genotype. The simplest model of evolution, whose equilibrium state is described by the quasispecies equation, is the Moran–Kingman model. For the sharp-peak landscape, we perform several exact computations and derive several exact formulas. We also obtain an exact formula for the quasispecies distribution, involving a series and the mean fitness. A very simple formula for the mean Hamming distance is derived, which is exact and does not require a specific asymptotic expansion (such as sending the length of the macromolecules to ∞ or the mutation probability to 0). With the help of these formulas, we present an original proof for the well-known phenomenon of the error threshold. We recover the limiting quasispecies distribution in the long-chain regime. We try also to extend these formulas to a general fitness landscape. We obtain an equation involving the covariance of the fitness and the Hamming class number in the quasispecies distribution. Going beyond the sharp-peak landscape, we consider fitness landscapes having finitely many peaks and a plateau-type landscape. Finally, within this framework, we prove rigorously the possible occurrence of the survival of the flattest, a phenomenon which was previously discovered by Wilke *et al.* (Nature **412**, 2001) and which has been investigated in several works (see e.g. Codoñer *et al.* (*PLOS Pathogens* **2**, 2006), Franklin *et al.* (*Artificial Life* **25**, 2019), Sardanyés *et al.* (*J. Theoret. Biol.* **250**, 2008), and Tejero *et al.* (*BMC Evolutionary Biol.* **11**, 2011)).

Keywords: Eigen’s model; mutation; selection; fitness; plateau

2010 Mathematics Subject Classification: Primary 92D15

Secondary 47A75, 37N25

1. Introduction

Some viruses are very hard to fight because of their genetic diversity: they mutate constantly and this allows them to escape the attacks of the immune system of their host. At the same time, the genotypes of an infecting population of viruses are strongly correlated; they look like a cloud of mutants centered around a specific genotype. This population structure, called a quasispecies, can be observed experimentally for viruses (*in vivo* studies have been conducted for HIV [17] and the hepatitis C virus [12]). Nowadays, deep sequencing techniques make it

Received 9 August 2023; accepted 15 August 2024.

* Postal address: 40 Rue René Cassin, 21850 Saint-Apollinaire, France. Email address: maximeberger74@gmail.com

** Postal address: Laboratoire de Mathématiques d’Orsay, Université Paris-Sud, CNRS, Université Paris-Saclay, 91405 Orsay, France. Email address: Raphael.Cerf@universite-paris-saclay.fr

© The Author(s), 2024. Published by Cambridge University Press on behalf of Applied Probability Trust.

possible to collect data on the structure of viral quasispecies. Yet the various biological mechanisms involved in the creation and evolution of quasispecies are far from being understood (see [8] for a recent review or the book [9] for a more comprehensive account). The design of simple mathematical models of quasispecies may be helpful in developing efficient antiviral therapies. In principle, a quasispecies can occur in any biological population whose evolution is driven by mutation and selection. In the next subsection, we present a very simple model for the evolution of a population under mutation and selection. We then write the equations describing the equilibrium of this model, thereby recovering directly the quasispecies equations associated with the equilibrium of Eigen's model. All the results presented afterwards concern the solutions of these equations.

1.1. The quasispecies equations

We consider a population of haploid individuals evolving under the conjugate effects of mutation and selection. We suppose that the population has reached an equilibrium around a specific well-adapted individual, called the wild type, and denoted by w^* . Individuals reproduce, yet the reproduction mechanism is error-prone, and mutations occur constantly. These mutations drive the genotypes away from w^* . Yet w^* has a selective advantage because it reproduces faster.

We would like to characterize this kind of equilibrium mathematically. We denote by E the set of possible genotypes, which we assume to be finite. Generic elements of E are denoted by the letters u, v . The Darwinian fitness of an individual having genotype u is denoted by $f(u)$; it can be thought of as its mean number of offspring. In addition, mutations occur in each reproduction cycle: an individual of type u might appear as the result of mutations from offspring of other types. Let us denote by $M(v, u)$ the probability that the offspring of an individual of type v is of type u . For simplicity, we will assume that all the coefficients of the mutation matrix M are positive. Of course, we have

$$\forall v \in E \quad \sum_{u \in E} M(v, u) = 1.$$

We introduce next the linear model, one of the simplest models for the evolution of a population with selection and mutation. We suppose that the successive generations do not overlap, and we denote by $N_n(u)$ the number of individuals of type u in generation n . The linear model assumes that an individual of type v produces a number of offspring proportional to its fitness $f(v)$, and that a proportion $M(v, u)$ of the offspring mutates and becomes of type u ; thus $N_{n+1}(u)$ is given by the formula

$$\forall u \in E \quad N_{n+1}(u) = \sum_{v \in E} N_n(v) f(v) M(v, u).$$

The trouble with this formula is that the sum is not necessarily an integer. A natural way to circumvent this problem is to develop stochastic population models, in such a way that the above formula describes the evolution of the mean number of individuals. The archetype of this kind of model is the Galton–Watson branching process. If we introduce in addition a constraint on the total size of the population, then we arrive at the classical Wright–Fisher model. Yet the randomness adds an extra layer of complexity; stochastic models are considerably harder to study. Another, simpler, possibility is to consider the proportion of each type of individual in the population, instead of the number, as Moran and Kingman did in the late seventies

[16, 18]. Let us denote by $x_n(u)$ the proportion of individuals of type u in generation n . The model proposed by Moran is given by

$$\forall u \in E \quad x_{n+1}(u) = \frac{\sum_{v \in E} x_n(v) f(v) M(v, u)}{\sum_{v \in E} x_n(v) f(v)}.$$

The denominator has been chosen to ensure that $\sum_{u \in E} x_{n+1}(u) = 1$. In fact, this model consists in iterating a deterministic map on the function encoding the proportion of each type present in the population. The possible equilibria of the model correspond to the fixed points of this map, that is, the solutions of the following equations:

$$\forall u \in E \quad x(u) \sum_{v \in E} x(v) f(v) = \sum_{v \in E} x(v) f(v) M(v, u) \quad (1.1)$$

subject to the constraint

$$\forall u \in E \quad x(u) \geq 0, \quad \sum_{u \in E} x(u) = 1. \quad (1.2)$$

We call these equations the quasispecies equations. They also characterize the equilibrium in the model originally developed by Eigen [10], which led to quasispecies theory. Some of our results are derived for the general quasispecies equations (1.1). To go further, we will focus on a particular choice of the set of genotypes E and of the mutation matrix M . For both practical and historical reasons, we make the same choice as Eigen and Schuster [11], which leads to the sharp-peak landscape.

Genotypes. We consider the different genotypes to be sequences of length $\ell \geq 1$ over an alphabet \mathcal{A} of cardinality $\kappa \geq 1$. Standard examples are $\mathcal{A} = \{0, 1\}$, $\kappa = 2$, $E = \{0, 1\}^\ell$, for binary sequences, and $\mathcal{A} = \{A, C, G, T\}$, $\kappa = 4$, $E = \{A, C, G, T\}^\ell$, to model DNA molecules. The i th component of a genotype u is denoted by $u(i)$. The Hamming distance d counts the number of differences between two sequences:

$$\forall u, v \in E \quad d(u, v) = \text{card} \{ 1 \leq i \leq \ell : u(i) \neq v(i) \}.$$

When we wish to work with the simplest model, we use binary sequences. When we seek to state a more general result, we use sequences over a finite alphabet.

Mutations. We suppose that mutations happen independently along the sequence with probability $q \in]0, 1[$, and that, upon mutation, a new letter of the alphabet is chosen with equal probability. For $u, v \in E$, the mutation probability $M(u, v)$ is given by

$$M(u, v) = \left(\frac{q}{\kappa - 1} \right)^{d(u, v)} (1 - q)^{\ell - d(u, v)}. \quad (1.3)$$

The final ingredient needed to completely specify the model is the fitness function. We will consider different types of fitness functions in the following subsections. For the initial models, for which the fitness function is quite simple, we work with the genotype space $E = \mathcal{A}^\ell$ where \mathcal{A} is a finite alphabet. For the more complicated fitness landscapes considered at the end, which contain a kind of plateau, we work with the genotype space of binary sequences $E = \{0, 1\}^\ell$.

1.2. The sharp-peak landscape

Ideally, we would like to have explicit formulas for x in terms of f and M . Unfortunately, there is little hope of obtaining such explicit formulas in the general case. So we consider the simplest non-neutral fitness function which comes to mind, the sharp-peak landscape: there is a privileged genotype, w^* , referred to as the wild type, which has a higher fitness than the rest. We work with the genotype space $E = \mathcal{A}^\ell$ where \mathcal{A} is a finite alphabet of cardinality κ . Let $\sigma > 1$, and let the fitness function f_{SP} be given by

$$\forall u \in E \quad f_{\text{SP}}(u) = \begin{cases} \sigma & \text{if } u = w^*, \\ 1 & \text{if } u \neq w^*. \end{cases} \quad (1.4)$$

This is the fitness function that Eigen studied in detail in his article [10]. Despite its apparent simplicity, this model leads to interesting mathematical questions, and it provides new insight into the genetic structure of a population. Eigen and Schuster [11] discussed the sharp-peak landscape (which is fully presented in Section 2.2) with the help of approximation techniques and in a specific asymptotic regime. We present here a new approach, which is more elementary. The sharp-peak landscape is studied in detail in Section 2. We perform several exact computations and derive several exact formulas. In particular, we obtain a remarkable equation for the mean fitness λ at equilibrium, which was previously discovered by Bratus *et al.*; see the formulas 5.5 and 5.6 in [3]. We state next a probabilistic version of this formula. Let S_ℓ be a random variable with distribution the binomial law with parameters ℓ and $(\kappa - 1)/\kappa$. Let \mathbb{E} denote the expectation; then the mean fitness λ at equilibrium satisfies

$$\frac{1}{\sigma - 1} = \mathbb{E} \left(\frac{1}{\frac{\lambda}{\rho^{S_\ell}} - 1} \right), \quad (1.5)$$

where we have set

$$\rho = 1 - \frac{\kappa q}{\kappa - 1}. \quad (1.6)$$

Bratus *et al.* exploited this equation to derive rigorous bounds on the mean fitness, but this formula has wider applications. Furthermore, in Section 2.4 we obtain an exact formula for the quasispecies distribution, involving a series and the mean fitness. A very simple formula for the mean Hamming distance is derived, which is exact and which does not require a specific asymptotic expansion, such as sending the length of the macromolecules to ∞ or the mutation probability to 0 (recall that the Hamming distance between two sequences is the number of sites where they differ). The mean Hamming distance \overline{Q} in the quasispecies distribution can be interpreted as the mean number of ones in a genotype at equilibrium, and the formula is

$$\overline{Q} = \frac{\ell q \lambda}{\lambda - \rho}, \quad (1.7)$$

where λ is the mean fitness of the population (here we assume that the wild type is the genotype $(0, \dots, 0)$ and the Hamming distances are computed with respect to the wild type).

We stress that all the formulas obtained in Section 2 are exact and hold for a fixed value of the length ℓ . In the next subsection and in Section 3, we perform asymptotic expansions of these formulas in the long-chain regime.

1.3. The error threshold

We shall demonstrate that the quasispecies equations for the sharp-peak landscape undergo a phenomenon similar to a phase transition. This is a classical and well-known result. However, we adopt an original approach for the proof; indeed, we will show that it follows easily from Equation (1.5). Furthermore, with the help of the exact formulas derived in Section 2.4, we recover the limiting quasispecies distribution in the long-chain regime in Section 3.3.

We consider the asymptotic regime, called the long-chain regime, where

$$\ell \rightarrow +\infty, \quad q \rightarrow 0, \quad \ell q \rightarrow a \in [0, +\infty]. \quad (1.8)$$

The parameter a is in fact the asymptotic mean number of observed mutations per individual in each reproduction cycle. We denote by λ the mean fitness at equilibrium and by $x(w^*)$ the proportion of the wild type w^* in the population at equilibrium.

Theorem 1. (Error threshold.) *For the sharp-peak landscape, we have the following dichotomy:*

- If $a > \ln \sigma$, then $\lambda \rightarrow 1$, whence $x(w^*) \rightarrow 0$.
- If $a < \ln \sigma$, then $\lambda \rightarrow \sigma e^{-a} > 1$, and

$$x(w^*) \rightarrow \frac{\sigma e^{-a} - 1}{\sigma - 1}.$$

Thus the limit of the proportion $x(w^*)$ of the wild type present in the population is positive only if $\sigma e^{-a} > 1$, or equivalently, if asymptotically $q < \ell^{-1} \ln \sigma$. The value

$$q^* = \frac{\ln \sigma}{\ell}$$

is the error threshold, described by Eigen as the critical mutation rate above which the wild type w^* disappears from the population. Whenever the occurrence of mutations is negligible compared to the sequence length, that is, for $a = 0$, we have $x(w^*) = 1$ and the only genotype present in the population is the wild type w^* . In the presence of a small number of mutations, that is, for $a > 0$, genetic diversity is constantly reintroduced in the population and there is a positive proportion of genotypes that differ from w^* . Yet most of the genotypes are very close to the wild type w^* . In fact, the population looks like a cloud of mutants centered around the wild type. As the mutation rate q is raised, the proportion of the wild type w^* present in the population decreases. When q reaches the error threshold, the wild type disappears from the population. More precisely, the proportion of the wild type becomes comparable to the proportions of the other genotypes. This is the error catastrophe: the selective advantage of the wild type w^* is annihilated by the mutations.

This kind of equilibrium was discovered within the framework of Eigen's model and was called a quasispecies [10, 11], as opposed to a species, which refers to a homogeneous solution in chemistry. In fact, Eigen's model is a model for the evolution of a population of macromolecules governed by a system of chemical reactions, and the laws of kinetics yield a differential system of equations whose equilibrium is described by the quasispecies equation. This system was historically analyzed with approximation and expansion techniques, which in the asymptotic regime (1.8) led to the discovery of the error threshold.

Although the original goal of Eigen was to understand the first stages of life on Earth, the notion of quasispecies and the error threshold had a profound impact on the understanding

of molecular evolution [8]. Indeed, many living organisms seem to satisfy approximately the scaling relation

$$\text{mutation rate} \sim \frac{1}{\text{genome length}}.$$

Unfortunately, for complex creatures, it is very difficult to estimate the mutation rate, which is usually extremely small. Viruses, however, have a rather high mutation rate, and the orders of magnitude of their genome length and their mutation rate are compatible with this scaling law. Moreover, some RNA viruses, such as HIV, evolve with a high mutation rate which seems to be close to an error threshold [8]. Why is that so?

In order to survive, a virus must achieve two goals. First, its genetic information must be preserved from one generation to another; hence its mutation rate has to be below the error threshold. Second, it has to escape the attacks of the immune system of its host, and to do so, it should mutate as fast as possible. The most efficient strategy is therefore to adjust the mutation rate to the left of the error threshold: this will achieve simultaneously a huge genetic variability, and the preservation of the pertinent genetic information across generations. Some promising antiviral strategies, called lethal mutagenesis, consist in using mutagenic drugs which slightly increase the mutation rate of the virus in order to induce an error catastrophe [1, 7, 8].

Most of the rigorous mathematical analysis of Eigen’s model deals with the sharp-peak landscape, or with landscapes presenting a lot of symmetries (see [2] for the analysis of a permutation-invariant fitness landscape and [4] for a review of the literature).

1.4. Extension to a general landscape

We try to extend the formulas obtained in Section 2 to a general fitness landscape. Here we work with the genotype space $E = \mathcal{A}^\ell$ where \mathcal{A} is a finite alphabet of cardinality κ . We make the following hypothesis on the fitness function f .

Hypothesis 1. *The fitness function f is larger than or equal to 1, and it is not identically equal to 1.*

In Theorem 2, we obtain an interesting equation involving the covariance of the fitness and the sum of an arbitrary function g over the letters of the sequences in the quasispecies distribution. This formula is a considerable generalization of the formula (1.7).

Theorem 2. *Let \mathcal{A} be a finite alphabet of cardinality κ and let $E = \mathcal{A}^\ell$. Let g be a function defined on \mathcal{A} with values in \mathbb{R} , and let $G : E \rightarrow \mathbb{R}$ be defined by*

$$\forall u \in E \quad G(u) = \sum_{i=1}^{\ell} g(u(i)).$$

For any fitness function f over the genotype set E which is larger than or equal to 1, and which is not identically equal to 1, we have

$$\bar{f}\bar{G} = \rho\bar{f}\bar{G} + \ell(1 - \rho)\bar{f}\bar{g}, \tag{1.9}$$

where ρ is given in (1.6) and

$$\begin{aligned} \bar{f}\bar{G} &= \sum_{u \in E} f(u) G(u) x(u), & \bar{g} &= \frac{1}{\kappa} \sum_{a \in \mathcal{A}} g(a), \\ \bar{f} &= \sum_{u \in E} f(u) x(u), & \bar{G} &= \sum_{u \in E} G(u) x(u). \end{aligned}$$

Let us look at the specific case of the sharp-peak landscape. We take $\kappa = 2$, $E = \{0, 1\}^\ell$, and the fitness function defined in the formula (1.4). For the function g , we take the identity function on $\{0, 1\}$, so that, for any $u \in E$, the value $G(u)$ is simply the Hamming class of u (that is, the number of ones present in u). In particular, we have $\overline{fG} = \overline{G}$ and $\overline{g} = 1/2$, so that the formula (1.9) becomes

$$\overline{fG} = \rho \overline{G} + \frac{\ell}{2}(1 - \rho)\overline{f}, \quad (1.10)$$

which is precisely the formula (1.7) that we obtained previously in this context. We can also consider the neutral case. If the fitness function is constant, then the quasispecies distribution is in fact the uniform distribution over the genotype space, so that $\overline{G} = \ell/2$, and the above relation (1.10) still holds. This is not too surprising, because it amounts to the case where the fitness σ of the wild type w^* is equal to 1.

The proof of Theorem 2 is given in Section 4, where we also perform several useful computations valid for a general state space E . The various formulas obtained in Section 4 shed new light on the classical phenomenon of the error threshold and the notion of quasispecies, which are discussed in Section 3.2. Finally, in Section 4, we outline a robust strategy for analyzing more general fitness landscapes. This strategy will be used to study the fitness landscapes introduced in the following subsections.

1.5. Finitely many peaks

Now that we have thoroughly understood the case of the sharp-peak landscape, we are ready to consider more complex landscapes. Here we work with the genotype space $E = \mathcal{A}^\ell$ where \mathcal{A} is a finite alphabet of cardinality κ . We make the following hypothesis on the fitness function f .

Hypothesis 2. *There exist a fixed integer k and k fixed values $\sigma_1, \dots, \sigma_k$ such that the following holds: for any values of ℓ, q , there exist k peaks w_1^*, \dots, w_k^* in E (whose locations may vary with ℓ and q) such that the fitness function f_{FP} is given by*

$$\forall u \in E \quad f_{FP}(u) = \begin{cases} \sigma_i & \text{if } u = w_i^*, \quad 1 \leq i \leq k, \\ 1 & \text{otherwise.} \end{cases} \quad (1.11)$$

We consider again the long-chain regime

$$\ell \rightarrow +\infty, \quad q \rightarrow 0, \quad \ell q \rightarrow a \in [0, +\infty].$$

To simplify the proofs, we suppose that $a < +\infty$. Indeed, the case $a = +\infty$ should be considered separately. We shall prove the following theorem.

Theorem 3. (Error threshold for finitely many peaks.) *Let us set*

$$\sigma = \max_{1 \leq i \leq k} \sigma_i.$$

Let $\lambda_{FP}(\ell, q)$ be the mean fitness for the model associated with the fitness function f_{FP} and the parameters ℓ, q . We have the following dichotomy:

- *If $a > \ln \sigma$, then $\lambda_{FP}(\ell, q) \rightarrow 1$.*
- *If $a < \ln \sigma$, then $\lambda_{FP}(\ell, q) \rightarrow \sigma e^{-a} > 1$.*

We see that the situation is very similar to the case of the sharp-peak landscape. The critical parameter depends now on the fitness of the highest peak. The conclusion is that, when there are finitely many peaks, the lower peaks do not really interfere with the highest peak. Two proofs of Theorem 3 are presented in Section 5.

1.6. Plateau

We have seen that a finite number of small peaks do not interact significantly with the highest peak. We consider here a fitness landscape which contains a kind of plateau, that is, a large subset of genotypes having a fitness $\sigma > 1$. More precisely, we consider the genotype space of binary sequences $E = \{0, 1\}^\ell$ and the usual mutation kernel M . To simplify the formulas, we suppose that the length ℓ is even. The Hamming class number $H(u)$ of $u \in \{0, 1\}^\ell$ is the number of ones present in u , that is,

$$H(u) = \sum_{1 \leq i \leq \ell} u(i).$$

We consider the fitness function f_{PL} defined as follows.

Hypothesis 3. *The fitness function f_{PL} is given by*

$$\forall u \in \{0, 1\}^\ell \quad f_{PL}(u) = \begin{cases} \sigma & \text{if } H(u) = \ell/2, \\ 1 & \text{otherwise.} \end{cases} \tag{1.12}$$

The plateau associated with the fitness function $f_{PL}(u)$ consists of the sequences having $\ell/2$ zeroes and $\ell/2$ ones. In particular, it contains sequences which are very far away in the Hamming distance. This might go against the intuitive notion of a plateau, especially in low-dimensional spaces, where we would usually take for a plateau a closed neighborhood around a fixed connected set. However, our goal is to find a subset of $\{0, 1\}^\ell$ which is very stable with respect to mutations, and the plateau of $f_{PL}(u)$ is the simplest example of such a set. We could also modify this set so that it is closer to our intuition of a plateau, by adding to it all the sequences which are at Hamming distance less than $\sqrt{\ell}$ from it. All the results presented here and in the subsequent section would hold for this larger set as well (but the proofs would become more complicated).

We consider again the long-chain regime with a finite:

$$\ell \rightarrow +\infty, \quad q \rightarrow 0, \quad \ell q \rightarrow a \in [0, +\infty[.$$

Theorem 4. *Let $\lambda_{PL}(\ell, q)$ be the mean fitness for the model associated with the fitness function f_{PL} and the parameters ℓ, q . For any $a > 0$, we have*

$$\lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} \lambda_{PL}(\ell, q) = \lambda(a, \sigma),$$

where $\lambda(a, \sigma)$ is the unique real number $\lambda > 1$ such that

$$\frac{1}{\sigma - 1} = \sum_{n \geq 1} \frac{e^{-an}}{\lambda^n} \sum_{k \geq 0} \frac{(an)^{2k}}{2^{2k}(k!)^2}. \tag{1.13}$$

The situation for the fitness landscape f_{PL} with the huge plateau is very different from the sharp-peak landscape. Indeed, the error threshold phenomenon described in Theorem 1 does

not occur, and there is no phase transition in the long-chain regime. It turns out that, if we shift the plateau away from $\ell/2$, then an error threshold phenomenon reappears. We consider the fitness function f_{PL} defined as follows.

Hypothesis 4. Let $\sigma > 1$ and $\alpha \in [0, 1]$. The fitness function f_{PL} is given by

$$\forall u \in \{0, 1\}^\ell \quad f_{PL}(u) = \begin{cases} \sigma & \text{if } H(u) = \lfloor \alpha \ell \rfloor, \\ 1 & \text{otherwise,} \end{cases} \quad (1.14)$$

where $\lfloor \alpha \ell \rfloor$ is the integer part of $\alpha \ell$.

The cases $\alpha = 0$ and $\alpha = 1$ correspond in fact to the sharp-peak landscape, which has already been studied. We shall therefore exclude these cases.

Theorem 5. Suppose that $\alpha \notin \{0, 1/2, 1\}$. Let $\lambda_{PL}(\ell, q)$ be the mean fitness for the model associated with the fitness function f_{PL} and the parameters ℓ, q . There exist two positive values $a_c^1(\sigma, \alpha) \leq a_c^2(\sigma, \alpha)$ such that the following holds: if $a \geq a_c^2(\sigma, \alpha)$, then

$$\lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} \lambda_{PL}(\ell, q) = 1.$$

If $a < a_c^1(\sigma, \alpha)$, then

$$\liminf_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} \lambda_{PL}(\ell, q) > 1.$$

Of course, it would be more satisfactory to know that $a_c^1(\sigma, \alpha) = a_c^2(\sigma, \alpha)$. This would be a consequence of the monotonicity of the function $\phi_\alpha(a)$ defined in (6.14); unfortunately this fact does not seem so obvious.

The proofs of Theorems 4 and 5 are presented in Section 6. The reason underlying these two results is the following. In the neutral fitness landscape, the most likely Hamming class is already the class $\ell/2$. By installing a plateau precisely on this Hamming class, we reinforce its stability. When we shift the plateau away from the Hamming class $\ell/2$, we create competition between two Hamming classes, and this leads to the occurrence of a phase transition.

1.7. Survival of the flattest

We finally prove rigorously within our framework the possible occurrence of the survival of the flattest, a phenomenon previously discovered by Wilke *et al.* [21], which has been investigated in several works (see e.g. [6, 14, 19, 20]). We now state the precise mathematical result we obtain. Here we will combine the sharp-peak landscape with a plateau. We consider the genotype space of binary sequences $E = \{0, 1\}^\ell$ and the usual mutation kernel M , with the following fitness function. To simplify the formulas, we suppose that the length ℓ is even.

Hypothesis 5. Let $\delta, \sigma \geq 1$ and let $f_{SP/PL}$ be the fitness function given by

$$\forall u \in \{0, 1\}^\ell \quad f_{SP/PL}(u) = \begin{cases} \delta & \text{if } u = 0 \cdots 0, \\ \sigma & \text{if } u \text{ contains } \frac{\ell}{2} \text{ zeroes and } \frac{\ell}{2} \text{ ones,} \\ 1 & \text{otherwise.} \end{cases} \quad (1.15)$$

Theorem 6. *Let $a > 0$, and let $\lambda(a, \sigma)$ be the unique real number $\lambda > 1$ such that*

$$\frac{1}{\sigma - 1} = \sum_{n \geq 1} \frac{e^{-an}}{\lambda^n} \sum_{k \geq 0} \frac{(an)^{2k}}{2^{2k}(k!)^2}. \tag{1.16}$$

Let $\lambda_{SP/PL}(\delta, \sigma, \ell, q)$ be the mean fitness for the model associated with the fitness function $f_{SP/PL}$ and the parameters ℓ, q . We have the following convergence in the long-chain regime:

$$\lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} \lambda_{SP/PL}(\delta, \sigma, \ell, q) = \max(\delta e^{-a}, \lambda(a, \sigma)).$$

Let us denote by $y(0)$ the fraction of the sequence $0 \cdots 0$ in the population at equilibrium, and by $y(\ell/2)$ the fraction of the sequences having $\ell/2$ zeroes and $\ell/2$ ones in the population at equilibrium.

- *If $\lambda(a, \sigma) > \delta e^{-a}$, then $y(0) \rightarrow 0, y(\ell/2) \rightarrow \frac{\lambda(a, \sigma) - 1}{\sigma - 1}$.*
- *If $\delta e^{-a} > \lambda(a, \sigma)$, then $y(\ell/2) \rightarrow 0, y(0) \rightarrow \frac{\delta e^{-a} - 1}{\delta - 1}$.*

A plateau is always more stable than a peak of the same height; that is, for any $a > 0$, any $\sigma > 1$, we have $\lambda(a, \sigma) > \sigma e^{-a}$. However, for a fixed value $a > 0$, there exist positive values $\sigma < \delta$ such that $\lambda(a, \sigma) > \delta e^{-a}$. Indeed, let us compute the right-hand side of Equation (1.16) with λ replaced by δe^{-a} :

$$\begin{aligned} \sum_{n \geq 1} \frac{e^{-an}}{(\delta e^{-a})^n} \sum_{k \geq 0} \frac{(an)^{2k}}{2^{2k}(k!)^2} &= \sum_{n \geq 1} \frac{1}{\delta^n} \sum_{k \geq 0} \frac{(an)^{2k}}{2^{2k}(k!)^2} \\ &\geq \sum_{n \geq 1} \frac{1}{\delta^n} \left(1 + \frac{(an)^2}{4} \right) = \frac{1}{\delta - 1} + \frac{a^2}{4} \frac{\delta(\delta + 1)}{(\delta - 1)^3}. \end{aligned}$$

So, for any value of $\delta > \sigma$ such that

$$\frac{1}{\sigma - 1} < \frac{1}{\delta - 1} + \frac{a^2}{4} \frac{\delta(\delta + 1)}{(\delta - 1)^3},$$

we will indeed have $\lambda(a, \sigma) > \delta e^{-a}$. For such values, the quasispecies will be located on the plateau, despite the fact that the height σ of the plateau is strictly less than the height of the peak. In this situation, we witness the survival of the flattest.

Theorem 6 is proved in Section 7. The route to proving Theorems 3 and 6 is quite long. It rests on the exact formulas proved in Section 4.2 and the general strategy developed in Section 4.

2. The sharp-peak landscape

Our first goal is to develop an exact formula for the quasispecies on the sharp-peak landscape. All our subsequent study of the sharp-peak landscape will rely on this formula.

2.1. A representation formula for the quasispecies

The computations performed in this subsection are in fact valid without any assumption on the geometry of the genotype space E . We merely suppose that E is finite and that

the coefficients of the matrix M are all positive. Suppose that $(x(u), u \in E)$ is a solution of the system (1.1) satisfying the constraint (1.2). Let λ be the mean fitness of this solution, defined by

$$\lambda = \sum_{v \in E} x(v)f(v) = \sigma x(w^*) + 1 - x(w^*). \quad (2.1)$$

Obviously, the mean fitness λ satisfies $1 \leq \lambda \leq \sigma$ and $\lambda = 1$ if and only if $x(w^*) = 0$. By hypothesis, all the coefficients of the mutation matrix M are positive, and it follows from Equation (1.1) that

$$\forall u \in E \quad x(u)\sigma \geq \sum_{v \in E} x(v)f(v)M(v, u) > 0. \quad (2.2)$$

Therefore $x(u) > 0$ for all $u \in E$. Equation (1.1) for a generic genotype u reads as follows:

$$\lambda x(u) = \sigma x(w^*)M(w^*, u) + \sum_{v \neq w^*} x(v)M(v, u). \quad (2.3)$$

We rewrite this equation as

$$x(u) = \frac{\sigma - 1}{\lambda} x(w^*)M(w^*, u) + \frac{1}{\lambda} \sum_{v \in E} x(v)M(v, u). \quad (2.4)$$

Using this very formula for $x(v)$, we replace $x(v)$ in the sum, thereby getting

$$\begin{aligned} x(u) = \frac{\sigma - 1}{\lambda} x(w^*)M(w^*, u) + \frac{\sigma - 1}{\lambda^2} \sum_{v \in E} x(w^*)M(w^*, v)M(v, u) \\ + \frac{1}{\lambda^2} \sum_{v, v' \in E} x(v')M(v', v)M(v, u). \end{aligned}$$

We iterate this process. After $N - 1$ steps, we get

$$\begin{aligned} x(u) = \sum_{1 \leq n \leq N} \frac{\sigma - 1}{\lambda^n} \sum_{v_1, \dots, v_{n-1} \in E} x(w^*)M(w^*, v_1) \cdots M(v_{n-1}, u) \\ + \frac{1}{\lambda^N} \sum_{v_1, \dots, v_N \in E} x(v_1)M(v_1, v_2) \cdots M(v_N, u). \end{aligned}$$

The sums of products involving the matrix M have a natural probabilistic interpretation. Let $(X_n)_{n \in \mathbb{N}}$ be the Markov chain on E with transition matrix M . The previous formula for $x(u)$ can now be rewritten as

$$\begin{aligned} x(u) = \sum_{1 \leq n \leq N} \frac{\sigma - 1}{\lambda^n} x(w^*) \mathbb{P}(X_n = u \mid X_0 = w^*) \\ + \frac{1}{\lambda^N} \sum_{v_1 \in E} x(v_1) \mathbb{P}(X_N = u \mid X_0 = v_1). \end{aligned}$$

By (2.1) and (2.2), the mean fitness λ is strictly larger than 1; therefore the last term goes to 0 as N goes to ∞ . Sending N to ∞ , and using the fact that

$$\lambda - 1 = (\sigma - 1)x(w^*), \quad (2.5)$$

we get the following result.

Proposition 1. Any solution $(x(u), u \in E)$ of the quasispecies equations (1.1) for the sharp-peak function f_{SP} defined in (1.4) satisfies

$$x(u) = \sum_{n \geq 1} \frac{\lambda - 1}{\lambda^n} \mathbb{P}(X_n = u \mid X_0 = w^*), \tag{2.6}$$

where $(X_n)_{n \geq 0}$ is the Markov chain on E with transition matrix M .

Let us take $u = w^*$ in the formula (2.6). Using (2.5) again, we obtain

$$\frac{1}{\sigma - 1} = \sum_{n \geq 1} \frac{1}{\lambda^n} \mathbb{P}(X_n = w^* \mid X_0 = w^*). \tag{2.7}$$

The right-hand side of this equation is a continuous decreasing function of λ , which is equal to $+\infty$ when $\lambda = 1$ and is less than or equal to $1/(\sigma - 1)$ when $\lambda = \sigma$. Therefore there exists exactly one value of λ in $[1, \sigma]$ which satisfies this equation. Our next goal is to obtain an explicit expression for $P(X_n = w^* \mid X_0 = w^*)$ and to perform the summation of the series appearing in (2.7). To do so, we shall make additional assumptions on the geometry of the fitness landscape.

2.2. Mutation dynamics

We now turn our attention to the genotype space $E = \mathcal{A}^\ell$ where \mathcal{A} is a finite alphabet of cardinality κ , and we consider the mutation scheme (1.3). In this model, the mutations occur independently at each site. An important consequence of this structural assumption is that the components $(X_n(i), 1 \leq i \leq \ell)$ of the Markov chain X_n are themselves independent Markov chains with state space \mathcal{A} and transition matrix

$$\begin{pmatrix} 1 - q & \frac{q}{\kappa - 1} & \cdots & \frac{q}{\kappa - 1} \\ \frac{q}{\kappa - 1} & 1 - q & \ddots & \frac{q}{\kappa - 1} \\ \vdots & \ddots & \ddots & \vdots \\ \frac{q}{\kappa - 1} & \cdots & \frac{q}{\kappa - 1} & 1 - q \end{pmatrix}.$$

The non-diagonal terms in this matrix are all equal to $q/(\kappa - 1)$. Since we want to compute $\mathbb{P}(X_n = w^* \mid X_0 = w^*)$, we shall lump together the letters which differ from the wild type w^* , and we shall do this for each component. More precisely, we define a process $(V_n)_{n \geq 0}$ by setting

$$\forall i \in \{1, \dots, \ell\} \quad V_n(i) = \begin{cases} 0 & \text{if } X_n(i) = w^*(i), \\ 1 & \text{if } X_n(i) \neq w^*(i). \end{cases}$$

The binary word V_n indicates the sites where X_n and w^* differ. In particular, V_n is a deterministic function of X_n . Thanks to the specific form of the transition matrix, it turns out that $V_n(i)$ is still a Markov chain. This is a particular case of the lumping theorem of Kemeny and Snell (see Theorem 6.3.2 of [15]). To see why, consider for instance the first two steps of the first

component $V_n(1)$, and let us make the following computation:

$$\begin{aligned} & \mathbb{P}(V_2(1) = 0, V_1(1) = 1 \mid V_0(1) = 0) \\ &= \sum_{x_1 \in \mathcal{A} \setminus \{w^*(1)\}} \mathbb{P}(X_2(1) = w^*(1), X_1(1) = x_1 \mid V_0(1) = 0) \\ &= \sum_{x_1 \in \mathcal{A} \setminus \{w^*(1)\}} \mathbb{P}(X_2(1) = w^*(1) \mid X_1(1) = x_1) \\ &\times \mathbb{P}(X_1(1) = x_1 \mid X_0(1) = w^*(1)) = (\kappa - 1) \times \frac{q}{\kappa - 1} \times \frac{q}{\kappa - 1} = q \frac{q}{\kappa - 1}. \end{aligned}$$

The simplification in the final result comes from the fact that the probabilities of finding or losing a letter from the wild type w^* are the same for all the letters in $\mathcal{A} \setminus \{w^*(1)\}$. A similar computation can be made for a finite number of steps. We conclude that $V_n(i)$ is the two-state Markov chain that we define and study next. Let $(E_n)_{n \geq 0}$ be the Markov chain with state space $\{0, 1\}$ and transition matrix

$$T = \begin{pmatrix} 1 - q & q \\ \frac{q}{\kappa - 1} & 1 - \frac{q}{\kappa - 1} \end{pmatrix}.$$

The eigenvalues of T are 1 and

$$\rho = 1 - \frac{\kappa q}{\kappa - 1}. \quad (2.8)$$

We will study some asymptotic regimes in which q is sent to 0, so we can suppose that $q \leq 1 - 1/\kappa$ (otherwise ρ might be negative, which would create additional complications). It is a standard exercise to compute the powers of T :

$$\forall n \geq 1 \quad T^n = \begin{pmatrix} \frac{1}{\kappa} + \frac{\kappa - 1}{\kappa} \rho^n & \frac{\kappa - 1}{\kappa} (1 - \rho^n) \\ \frac{1}{\kappa} (1 - \rho^n) & \frac{\kappa - 1}{\kappa} + \frac{1}{\kappa} \rho^n \end{pmatrix}.$$

Here is a simple, illuminating way to realize the Markov chain $(E_n)_{n \geq 0}$ and to understand the expression for the n th power T^n . Let $(\varepsilon_n)_{n \geq 1}$ be a sequence of independent and identically distributed Bernoulli random variables with parameter ρ . Suppose that $E_{n-1} = e \in \{0, 1\}$. If $\varepsilon_n = 1$, then we set $E_n = e$. If $\varepsilon_n = 0$, then we choose a letter uniformly over \mathcal{A} , independently of the past history up to time n , and we set $E_n = 0$ if the chosen letter is the one of the wild type and $E_n = 1$ otherwise. Now, the event $E_n = E_0$ can occur in two different ways. Either $\varepsilon_1 = \dots = \varepsilon_n = 1$, or one of $\varepsilon_1, \dots, \varepsilon_n$ is zero, in which case $E_n = 0$ with probability $1/\kappa$ and $E_n = 1$ with probability $(\kappa - 1)/\kappa$; thus

$$\begin{aligned} \mathbb{P}(E_n = 0 \mid E_0 = 0) &= \rho^n + (1 - \rho^n) \frac{1}{\kappa}, \\ \mathbb{P}(E_n = 1 \mid E_0 = 1) &= \rho^n + (1 - \rho^n) \frac{\kappa - 1}{\kappa}, \end{aligned}$$

and we recover the expression for the diagonal coefficients of T^n . In words, the status E_n at step n is the same as at time 0 if no mutation has occurred, or if the last mutation results in the same letter as the wild type (case $E_0 = 0$) or in a different letter (case $E_0 = 1$). Similarly, the

event $E_n \neq E_0$ can occur only if one of $\varepsilon_1, \dots, \varepsilon_n$ is zero, and the last mutation event yields the adequate letter; thus

$$\begin{aligned} \mathbb{P}(E_n = 1 \mid E_0 = 0) &= (1 - \rho^n) \frac{\kappa - 1}{\kappa}, \\ \mathbb{P}(E_n = 0 \mid E_0 = 1) &= (1 - \rho^n) \frac{1}{\kappa}. \end{aligned}$$

Now the probability $\mathbb{P}(X_n = w^* \mid X_0 = w^*)$ can be rewritten with the help of V_n and E_n as

$$\begin{aligned} \mathbb{P}(X_n = w^* \mid X_0 = w^*) &= \mathbb{P}(X_n(i) = w^*(i), 1 \leq i \leq \ell \mid X_0 = w^*) \\ &= \mathbb{P}(V_n(i) = 0, 1 \leq i \leq \ell \mid X_0 = w^*) \\ &= \prod_{1 \leq i \leq \ell} \mathbb{P}(V_n(i) = 0 \mid X_0 = w^*) \\ &= \left(\mathbb{P}(E_n = 0 \mid E_0 = 0) \right)^\ell = \left(\frac{1}{\kappa} + \frac{\kappa - 1}{\kappa} \rho^n \right)^\ell. \end{aligned} \tag{2.9}$$

2.3. The equation for the mean fitness

We now have all the material necessary to obtain Equation (1.5). Substituting the expression obtained in (2.9) into the formula (2.7), we get

$$\frac{1}{\sigma - 1} = \sum_{n \geq 1} \frac{1}{\lambda^n} \left(\frac{1}{\kappa} + \frac{\kappa - 1}{\kappa} \rho^n \right)^\ell. \tag{2.10}$$

We expand the power in ℓ , exchange the summations, and re-sum the series as follows:

$$\begin{aligned} \frac{1}{\sigma - 1} &= \sum_{n \geq 1} \frac{1}{\lambda^n} \sum_{0 \leq k \leq \ell} \binom{\ell}{k} \left(\frac{1}{\kappa} \right)^{\ell - k} \left(\frac{\kappa - 1}{\kappa} \rho^n \right)^k \\ &= \frac{1}{\kappa^\ell} \sum_{0 \leq k \leq \ell} \binom{\ell}{k} (\kappa - 1)^k \sum_{n \geq 1} \frac{\rho^{kn}}{\lambda^n} \\ &= \frac{1}{\kappa^\ell} \sum_{0 \leq k \leq \ell} \binom{\ell}{k} (\kappa - 1)^k \frac{1}{\frac{\lambda}{\rho^k} - 1}. \end{aligned}$$

Thus the equation satisfied by the mean fitness λ reads

$$\frac{1}{\sigma - 1} = \frac{1}{\kappa^\ell} \sum_{0 \leq k \leq \ell} \binom{\ell}{k} (\kappa - 1)^k \frac{1}{\frac{\lambda}{\rho^k} - 1}. \tag{2.11}$$

This equation was discovered by Bratus *et al.* (see the formulas (5.5) and (5.6) in [3]), who exploited it to derive rigorous bounds on the mean fitness. This formula calls naturally for a probabilistic interpretation. Let S_ℓ be a random variable with distribution the binomial law with parameters ℓ and $(\kappa - 1)/\kappa$. Denoting the expectation by \mathbb{E} , we can rewrite Equation (2.11) as

$$\frac{1}{\sigma - 1} = \mathbb{E} \left(\frac{1}{\frac{\lambda}{\rho^{S_\ell}} - 1} \right). \tag{2.12}$$

Our analysis of the quasispecies equations on the sharp-peak landscape rests on the identity (2.10) or its equivalent form (2.12). Indeed, if we manage to estimate the mean fitness λ , then we will also have an estimate for the proportion $x(w^*)$ of the wild type present in the population. Moreover, once we know λ or $x(w^*)$, the proportions of the other types are completely determined by the formula (2.6).

2.4. An exact formula for the quasispecies

Loosely speaking, a quasispecies is a cloud of mutants centered around a specific genotype. The structure of this cloud depends on the parameters σ and q . In fact, the proportion of the wild type w^* decreases as the mutation rate q increases, and it becomes comparable to the proportions of the other genotypes when q reaches the error threshold q^* . This fascinating phenomenon will be established rigorously in Section 3.2. When q is a little below q^* , we observe a cloud of mutants centered around w^* which contains a very small proportion of the wild type w^* ; the vast majority of the genotypes present in the population will differ from w^* . Yet the genetic information carried out by w^* is still present in the population and determines its structure. This paradoxical situation has led several biologists to argue that selection operates at the level of the quasispecies, and not at the level of individuals [8]. Thus an important goal is to understand better the statistical composition of the cloud of mutants. In order to do so, we classify the genotypes according to their Hamming distance from the wild type w^* . For $k \geq 1$, we define the Hamming class k as the set \mathcal{H}_k of the genotypes which differ from w^* at exactly k indices, i.e.,

$$\mathcal{H}_k = \{ u \in E : d(u, w^*) = k \}. \quad (2.13)$$

We shall exploit further the formula (2.6) in order to derive an exact formula for the proportion $\mathcal{Q}(k)$ of the genotypes belonging to the Hamming class k . For $k \geq 1$, we have, thanks to the formula (2.6),

$$\begin{aligned} \mathcal{Q}(k) &= \sum_{u \in \mathcal{H}_k} \sum_{n \geq 1} \frac{\lambda - 1}{\lambda^n} \mathbb{P}(X_n = u \mid X_0 = w^*) \\ &= \sum_{n \geq 1} \frac{\lambda - 1}{\lambda^n} \sum_{u \in \mathcal{H}_k} \mathbb{P}(X_n = u \mid X_0 = w^*) \\ &= \sum_{n \geq 1} \frac{\lambda - 1}{\lambda^n} \mathbb{P}(X_n \in \mathcal{H}_k \mid X_0 = w^*). \end{aligned} \quad (2.14)$$

We have already noticed that the components of X_n , $(X_n(i), 1 \leq i \leq \ell)$, are independent Markov chains. Therefore, using the notation of Section 2.2, we have

$$\begin{aligned} \mathbb{P}(X_n \in \mathcal{H}_k \mid X_0 = w^*) &= \mathbb{P}\left(\sum_{1 \leq i \leq \ell} V_n(i) = k \mid X_0 = w^* \right) \\ &= \binom{\ell}{k} \left(\mathbb{P}(E_n = 1 \mid E_0 = 0) \right)^k \left(\mathbb{P}(E_n = 0 \mid E_0 = 0) \right)^{\ell - k} \\ &= \binom{\ell}{k} \left((1 - \rho^n) \frac{\kappa - 1}{\kappa} \right)^k \left(\rho^n + (1 - \rho^n) \frac{1}{\kappa} \right)^{\ell - k}. \end{aligned} \quad (2.15)$$

Recall that $\rho = 1 - \kappa q / (\kappa - 1)$. Plugging (2.15) into (2.14), we get an exact formula for the quasispecies distribution, in terms of a series involving λ and ρ :

$$Q(k) = \sum_{n \geq 1} \frac{\lambda - 1}{\lambda^n} \binom{\ell}{k} \left((1 - \rho^n)^{\frac{\kappa - 1}{\kappa}} \right)^k \left(\rho^n + (1 - \rho^n) \frac{1}{\kappa} \right)^{\ell - k}. \tag{2.16}$$

We can even compute the sum of the series, by developing the powers, and obtain a finite algebraic formula:

$$Q(k) = (\lambda - 1) \binom{\ell}{k} \frac{(\kappa - 1)^k}{\kappa^\ell} \sum_{i=0}^k \sum_{j=0}^{\ell - k} \binom{k}{i} \binom{\ell - k}{j} \frac{(-1)^i (\kappa - 1)^j \rho^{i+j}}{\lambda - \rho^{i+j}}. \tag{2.17}$$

Of course this formula is complicated, yet it expresses completely the dependence of $Q(k)$ on λ and q , and it shows the complexity of the sharp-peak landscape. Surprisingly, the mean and the variance of the Hamming distance of the quasispecies distribution have very simple expressions. Indeed, the mean of the Hamming distance of the distribution \bar{Q} is given by

$$\bar{Q} = \sum_{k=0}^{\ell} k Q(k) = \frac{\ell q \lambda}{\lambda - \rho}. \tag{2.18}$$

To compute this mean, we rely on the formula (2.16) to write

$$\begin{aligned} \bar{Q} &= \sum_{k=0}^{\ell} k Q(k) \\ &= \sum_{n \geq 1} \frac{\lambda - 1}{\lambda^n} \sum_{k=0}^{\ell} k \binom{\ell}{k} \left((1 - \rho^n)^{\frac{\kappa - 1}{\kappa}} \right)^k \left(\rho^n + (1 - \rho^n) \frac{1}{\kappa} \right)^{\ell - k}. \end{aligned} \tag{2.19}$$

In the inner sum we recognize the expectation of a binomial law, whence

$$\bar{Q} = \sum_{n \geq 1} \frac{\lambda - 1}{\lambda^n} \ell (1 - \rho^n)^{\frac{\kappa - 1}{\kappa}}.$$

By summing the geometric series and using the definition of ρ , we obtain the formula (2.18). In fact, all the successive moments of the quasispecies distribution can be computed in this way.

3. The error threshold

We continue to consider the case of the sharp-peak landscape described in Section 1.2. In Section 2, we performed exact computations with the length ℓ fixed. Here we shall perform asymptotic expansions in the long-chain regime. We will revisit the classical error threshold phenomenon and compute the limiting quasispecies distribution.

3.1. An upper bound on λ

The goal of this subsection is to prepare the ground for the technical proof of the error threshold result. We start from the identity (2.12), and we wish to obtain an upper bound

on λ . From the weak law of large numbers, we know that the most likely values for S_ℓ are those around its mean $(\kappa - 1)\ell/\kappa$, so we pick a positive number ε and split the expectation as follows:

$$\begin{aligned} \frac{1}{\sigma - 1} &= \sum_{0 \leq k \leq \ell} \mathbb{P}(S_\ell = k) \frac{1}{\frac{\lambda}{\rho^k} - 1} \\ &= \sum_{k: |k - (\kappa - 1)\ell/\kappa| > \varepsilon \ell} \mathbb{P}(S_\ell = k) \frac{1}{\frac{\lambda}{\rho^k} - 1} + \sum_{k: |k - (\kappa - 1)\ell/\kappa| \leq \varepsilon \ell} \mathbb{P}(S_\ell = k) \frac{1}{\frac{\lambda}{\rho^k} - 1}. \end{aligned} \quad (3.1)$$

Taking advantage of the fact that $\rho \leq 1$, we bound ρ^k by 1 in the first sum and by $\rho^{(\kappa - 1)\ell/\kappa - \varepsilon \ell}$ in the second; this leads to the inequality

$$\frac{1}{\sigma - 1} \leq \frac{1}{\lambda - 1} \mathbb{P}\left(\left|S_\ell - \frac{\kappa - 1}{\kappa} \ell\right| > \varepsilon \ell\right) + \frac{1}{\frac{\lambda}{\rho^{(\kappa - 1)\ell/\kappa - \varepsilon \ell}} - 1}. \quad (3.2)$$

We apply the Chebyshev inequality to the binomial random variable S_ℓ to obtain the classical estimate

$$\mathbb{P}\left(\left|S_\ell - \frac{\kappa - 1}{\kappa} \ell\right| > \varepsilon \ell\right) \leq \frac{\text{variance}(S_\ell)}{\varepsilon^2 \ell^2} = \frac{\kappa - 1}{\kappa^2 \varepsilon^2 \ell}. \quad (3.3)$$

Plugging the Chebyshev inequality into (3.2), we conclude that

$$\forall \varepsilon > 0 \quad \frac{1}{\sigma - 1} \leq \frac{\kappa - 1}{\kappa^2 \varepsilon^2 \ell (\lambda - 1)} + \frac{1}{\frac{\lambda}{\rho^{(\kappa - 1)\ell/\kappa - \varepsilon \ell}} - 1}. \quad (3.4)$$

Let us pause for one moment to look at this inequality. Obviously, as λ goes to ∞ , the right-hand side goes to 0, and the inequality cannot hold. In other words, for this inequality to hold, λ must not be too large. Thus, from this inequality, we should be able to derive an upper bound on λ . To get the best possible upper bound, we could rewrite it as a polynomial inequality of degree two in λ and compute the associated roots, but this leads to messy expressions. Instead, after a few trials (hidden from the reader), we obtain the following inequality.

Lemma 1. *The mean fitness λ satisfies*

$$\lambda \leq \max\left(1 + \frac{4\sigma}{\ell^{1/4}}, \sigma\left(1 - \frac{\kappa q}{\kappa - 1}\right)^{(\kappa - 1)\ell/\kappa - \ell^{2/3}} \left(1 + \frac{1}{\ell^{1/13}}\right)\right). \quad (3.5)$$

Proof. Suppose that λ is larger than the upper bound in (3.5). We check that this is not compatible with the inequality (3.4) with the choice $\varepsilon = \ell^{-1/3}$. Indeed, we have on the one hand

$$\lambda > 1 + \frac{4\sigma}{\ell^{1/4}} \quad \implies \quad \frac{1}{\ell^{1/3}(\lambda - 1)} \leq \frac{1}{4\sigma \ell^{1/12}}.$$

On the other hand,

$$\begin{aligned} \lambda > \sigma\left(1 - \frac{\kappa q}{\kappa - 1}\right)^{(\kappa - 1)\ell/\kappa - \ell^{2/3}} \left(1 + \frac{1}{\ell^{1/13}}\right) &\implies \\ \frac{1}{\frac{\lambda}{\rho^{(\kappa - 1)\ell/\kappa - \ell^{2/3}}} - 1} &\leq \frac{1}{\sigma\left(1 + \frac{1}{\ell^{1/13}}\right) - 1}. \end{aligned}$$

Applying the mean value theorem to the function $x \mapsto 1/(\sigma - 1 + x)$ on the interval $[0, \sigma/\ell^{1/13}]$, we have

$$\frac{1}{\sigma\left(1 + \frac{1}{\ell^{1/13}}\right) - 1} \leq \frac{1}{\sigma - 1} - \frac{1}{(2\sigma - 1)^2} \frac{\sigma}{\ell^{1/13}} \leq \frac{1}{\sigma - 1} - \frac{1}{4\sigma\ell^{1/13}}. \tag{3.6}$$

Combining the previous inequalities, we obtain that

$$\frac{\kappa - 1}{\kappa^2\ell^{1/3}(\lambda - 1)} + \frac{1}{\frac{\lambda}{\rho^{(\kappa-1)\ell/\kappa - \ell^{2/3}} - 1}} \leq \frac{1}{\sigma - 1} + \frac{1}{4\sigma} \left(\frac{1}{\ell^{1/12}} - \frac{1}{\ell^{1/13}} \right) < \frac{1}{\sigma - 1},$$

which stands in contradiction to the inequality (3.4) with $\varepsilon = \ell^{-1/3}$. □

3.2. Revisiting the error threshold

Because of its importance, we shall give two proofs of Theorem 1. The first proof is a soft proof based on a compactness argument, whose starting point is the initial quasispecies equation. The second proof is a more technical proof, whose starting point is the Equation (2.10) discovered by Bratus *et al.* Naturally, the second proof is more informative: it yields a control on the speed of convergence and opens the way to performing an asymptotic expansion of λ with respect to ℓ and q .

From Equation (2.1), we can express $x(w^*)$ as

$$x(w^*) = \frac{\lambda - 1}{\sigma - 1}. \tag{3.7}$$

Therefore it is enough to study the asymptotic behavior of λ . We shall prove that

$$\lambda \longrightarrow \max(1, \sigma e^{-a}). \tag{3.8}$$

Soft proof. In the long-chain regime, we have

$$\sigma(1 - q)^\ell \longrightarrow \sigma e^{-a}. \tag{3.9}$$

The quasispecies equation (1.1) associated with w^* reads

$$\lambda x(w^*) = \sigma x(w^*)M(w^*, w^*) + \sum_{v \in E \setminus \{w^*\}} x(v)M(v, w^*). \tag{3.10}$$

We know that $x(w^*) > 0$; therefore we deduce from (3.10) that

$$\lambda \geq \sigma M(w^*, w^*) = \sigma(1 - q)^\ell. \tag{3.11}$$

Using the lower bound (3.11) and the fact that $\lambda \geq 1$, we see that

$$\liminf \lambda \geq \max(1, \sigma e^{-a}). \tag{3.12}$$

Suppose that λ does not converge towards $\max(1, \sigma e^{-a})$. In this case, there exist $\varepsilon > 0$ and two sequences of parameters $(\ell_n)_{n \geq 0}, (q_n)_{n \geq 0}$ such that

$$\begin{aligned} \ell_n &\rightarrow +\infty, & q_n &\rightarrow 0, & \ell_n q_n &\rightarrow a, \\ \forall n \geq 0 & \quad \lambda(\ell_n, q_n) &\geq \max(1, \sigma e^{-a}) + \varepsilon. \end{aligned} \quad (3.13)$$

Using (3.7) and (3.13), we obtain a lower bound on $x(w^*)$:

$$x(w^*) \geq \frac{\varepsilon}{\sigma - 1}. \quad (3.14)$$

We bound from above the quasispecies equation (3.10) associated with w^* , as follows:

$$\lambda x(w^*) \leq \sigma x(w^*)(1 - q)^\ell + (1 - x(w^*))q, \quad (3.15)$$

where we have used the fact that $M(v, w^*) \leq q$ whenever $v \neq w^*$. Using (3.14) and (3.15) together, we obtain that

$$\forall n \geq 0 \quad \lambda(\ell_n, q_n) \leq \sigma(1 - q_n)^{\ell_n} + \frac{\sigma - 1}{\varepsilon} q_n.$$

Sending n to ∞ , we see that this inequality is not coherent with the inequality (3.13). Therefore it must be the case that λ converges towards $\max(1, \sigma e^{-a})$.

Technical proof. This proof is entirely based on the equation (2.10) discovered by Bratus *et al.* (see the formulas 5.5 and 5.6 in [3]) and its probabilistic interpretation (2.12). Let us start with a simple yet interesting inequality on λ . The map

$$\phi : x \in [0, +\infty[\mapsto \frac{1}{\frac{\lambda}{\rho^x} - 1}$$

is convex (recall that we suppose that $\rho > 0$). The equation (2.10) (or rather (2.12)) can be rewritten with the help of the function ϕ as

$$\frac{1}{\sigma - 1} = \mathbb{E}(\phi(S_\ell)), \quad (3.16)$$

where S_ℓ is a random variable with distribution the binomial law with parameters ℓ and $(\kappa - 1)/\kappa$. By the classical Jensen inequality, we therefore have

$$\frac{1}{\sigma - 1} \geq \phi(\mathbb{E}(S_\ell)) = \phi\left(\frac{\kappa - 1}{\kappa} \ell\right). \quad (3.17)$$

This yields the inequality

$$\lambda \geq \sigma \rho^{(\kappa-1)\ell/\kappa} = \sigma \left(1 - \frac{\kappa q}{\kappa - 1}\right)^{(\kappa-1)\ell/\kappa}. \quad (3.18)$$

Notice that this inequality is not better than the lower bound (3.11) used in the soft proof. However, it can certainly be improved with additional work. In the long-chain regime, we have

$$\sigma \left(1 - \frac{\kappa q}{\kappa - 1}\right)^{(\kappa-1)\ell/\kappa} \longrightarrow \sigma e^{-a}. \quad (3.19)$$

We combine the lower bound (3.18) and the upper bound (3.5) (which were both derived from the equation (2.10)) with the limit stated in (3.19) to obtain the desired conclusion.

3.3. The limiting quasispecies distribution

In Section 2.4, we computed an exact expression for the mean Hamming distance under the quasispecies distribution, which was

$$\bar{Q} = \sum_{k=0}^{\ell} k \mathcal{Q}(\sigma, a)(k) = \frac{\ell q \lambda}{\lambda - \rho}. \quad (3.20)$$

This formula is particularly illuminating for the error threshold phenomenon. We consider again the long-chain regime

$$\ell \rightarrow +\infty, \quad q \rightarrow 0, \quad \ell q \rightarrow a \in [0, +\infty]. \quad (3.21)$$

We have the following dichotomy:

- If $a > \ln \sigma$, then $\lambda \rightarrow 1$, $\lambda - \rho \rightarrow 0$, whence $\bar{Q} \rightarrow +\infty$.
- If $a < \ln \sigma$, then $\lambda \rightarrow \sigma e^{-a} > 1$, and

$$\bar{Q} \rightarrow \frac{a \sigma e^{-a}}{\sigma e^{-a} - 1}. \quad (3.22)$$

We can go even further; indeed we can perform the expansion of the formula (2.15) in the regime (3.21). From this we obtain

$$\mathbb{P}(X_n \in \mathcal{H}_k \mid X_0 = w^*) \sim \frac{\ell^k}{k!} (nq)^k (1 - nq)^\ell \sim \frac{(na)^k}{k!} e^{-na};$$

thus the asymptotic distribution of X_n is the Poisson distribution with parameter na . Combined with the formula (3.8), this yields that, for any fixed $n \geq 1$,

$$\frac{1}{\lambda^n} \mathbb{P}(X_n \in \mathcal{H}_k \mid X_0 = w^*) \sim \frac{(na)^k}{\sigma^n k!}.$$

Through brute substitution of this expansion into the formula (2.14), and again applying (3.8), we finally get

$$\mathcal{Q}(k) \sim (\sigma e^{-a} - 1) \frac{a^k}{k!} \sum_{n \geq 1} \frac{n^k}{\sigma^n}.$$

Some steps in this computation must be made rigorous, but this can be done by a routine application of the dominated convergence theorem. In fact, this formula was obtained in [5] while studying the quasispecies arising in a stochastic model for the evolution of a finite population. Thanks to the formula, we can easily plot the quasispecies distribution (Figures 1 and 2) and study its dependence on the parameters σ , a . In the quasispecies literature, this was previously done by numerically integrating the differential system (see for instance [9, Figure 3, p. 9]). In addition, we can now perform simple theoretical computations with the help of this distribution. For instance, the mutation rate at which the proportion of the Hamming class 1 becomes larger than the proportion of the wild type w^* is $a = 1 - 1/\sigma$.

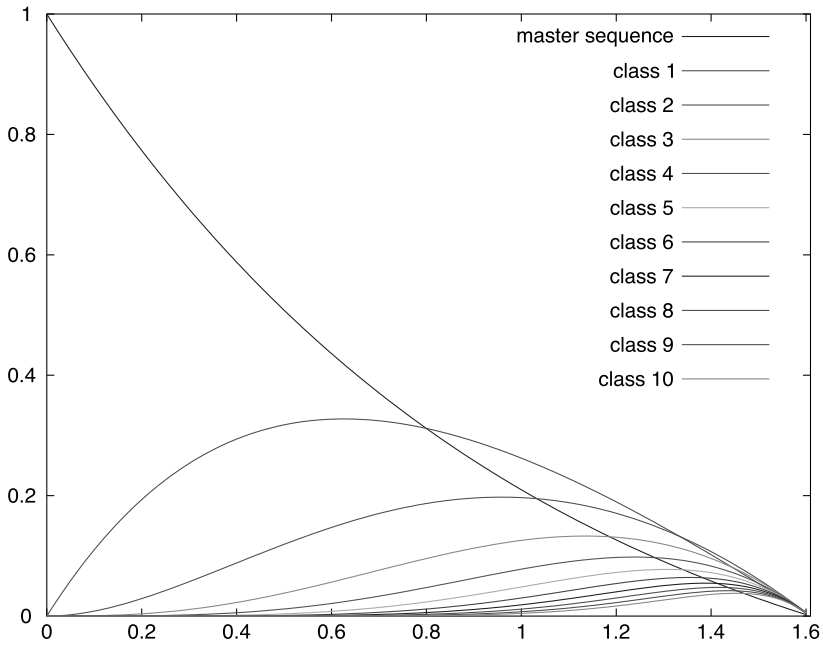


FIGURE 1. Fractions of the different types as functions of a for $\sigma = 5$.

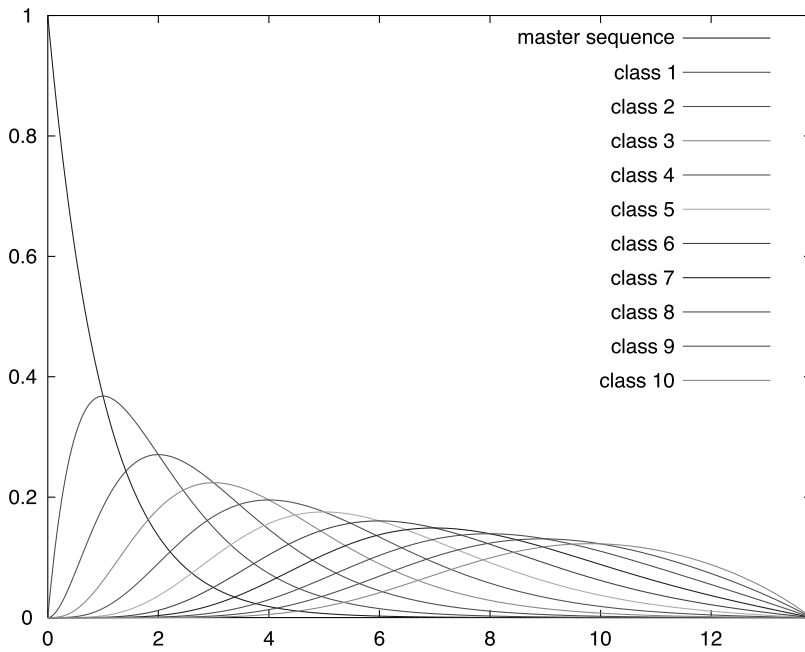


FIGURE 2. Fractions of the different types as functions of a for $\sigma = 10^6$.

4. Extension to a general landscape

Our starting point is the quasispecies equation (1.1), rewritten in the following form:

$$\forall u \in E \quad \lambda x(u) = \sum_{v \in E} x(v) f(v) M(v, u) \quad (4.1)$$

subject to the constraint

$$\forall u \in E \quad x(u) \geq 0, \quad \sum_{u \in E} x(u) = 1, \quad (4.2)$$

where

$$\lambda = \sum_{v \in E} x(v) f(v)$$

is the mean fitness, or equivalently the Perron–Frobenius eigenvalue of the matrix $(f(v)M(v, u))_{u, v \in E}$.

4.1. Reformulation of the quasispecies equations

The computations performed in this subsection are in fact valid without any assumption on the geometry of the genotype space E . Thanks to the hypothesis on the fitness function, we have that $\lambda \geq 1$. Moreover, since the matrix M has only positive entries and since the fitness function is not identically equal to 1, the Perron–Frobenius eigenvalue is strictly larger than 1, i.e., $\lambda > 1$. We rewrite the equation (4.1) as

$$\forall u \in E \quad \lambda x(u) - \sum_{v \in E} x(v) M(v, u) = \sum_{v \in E} x(v) (f(v) - 1) M(v, u). \quad (4.3)$$

We denote by x^t the row vector which is the transpose of the column vector x , and we introduce the square matrix whose diagonal coefficients are the fitness values $(f(u), u \in E)$:

$$F = \begin{pmatrix} \ddots & 0 & 0 \\ 0 & f(u) & 0 \\ 0 & 0 & \ddots \end{pmatrix}. \quad (4.4)$$

After dividing by λ (recall that $\lambda > 1$), we can rewrite Equation (4.3) in matrix form as

$$x^t \left(I - \frac{1}{\lambda} M \right) = x^t \frac{1}{\lambda} (F - I) M. \quad (4.5)$$

Now the matrix M is a stochastic matrix, so its spectral radius is equal to 1. Since $\lambda > 1$, the matrix $I - \frac{1}{\lambda} M$ is invertible, and its inverse is given by the geometric series

$$\left(I - \frac{M}{\lambda} \right)^{-1} = \sum_{n \geq 0} \frac{M^n}{\lambda^n}.$$

Plugging this identity into Equation (4.5), we obtain

$$x^t = x^t (F - I) \sum_{n \geq 1} \frac{M^n}{\lambda^n}. \quad (4.6)$$

Let us introduce the set W^* of the genotypes where the fitness function is strictly larger than 1:

$$W^* = \{u \in E : f(u) > 1\}.$$

We call the set W^* the set of wild types. We define the matrix N_λ , indexed by the elements of W^* , as follows:

$$\forall v, w \in W^* \quad N_\lambda(v, w) = \left((F - I) \sum_{n \geq 1} \frac{M^n}{\lambda^n} \right)(v, w). \quad (4.7)$$

With this notation, the linear system (4.6) of size κ^ℓ can be split into the system of size $|W^*|$ given by

$$\forall v \in W^* \quad x(v) = \sum_{w \in W^*} x(w) N_\lambda(w, v) \quad (4.8)$$

and the $\kappa^\ell - |W^*|$ remaining equations

$$\forall v \in E \setminus W^* \quad x(v) = \sum_{w \in W^*} x(w)(f(w) - 1) \sum_{n \geq 1} \frac{M^n}{\lambda^n}(w, v). \quad (4.9)$$

The system (4.8) expresses that the vector $(x(v), v \in W^*)$ is a left Perron–Frobenius eigenvector of the matrix N_λ , and that the Perron–Frobenius eigenvalue of N_λ is equal to 1. The equations (4.9) show that the remaining coordinates $(x(v), v \in E \setminus W^*)$ are completely determined by $(x(v), v \in W^*)$. These considerations suggest that the quasispecies equation can in principle be solved through the following procedure.

General strategy. We start by isolating the set W^* of the wild types, and we form the matrix $(N_\lambda(w, v))_{v, w \in W^*}$, defined in the formula (4.7), where $\lambda > 1$ is considered as a parameter. Now, the Perron–Frobenius eigenvalue of N_λ is a decreasing function of λ which tends to 0 as λ goes to $+\infty$ and to $+\infty$ as λ goes to 1. We choose for λ the unique value such that the Perron–Frobenius eigenvalue of N_λ is equal to 1. Once this value is fixed, we solve the system (4.8) on W^* . The remaining coordinates of x on $E \setminus W^*$ are determined by the equations (4.9).

A general lower bound on λ . We close this section with a general lower bound on λ . The quasispecies equations imply that

$$\forall u \in E \quad \lambda x(u) \geq x(u) f(u) M(u, u) = x(u) f(u) (1 - q)^\ell.$$

Since $x(u) > 0$ for any $u \in E$, this readily implies that

$$\lambda \geq \left(\max_E f \right) (1 - q)^\ell. \quad (4.10)$$

4.2. Computation of $\sum_{n \geq 1} M^n / \lambda^n$

From now on, we work with the genotype space $E = \mathcal{A}^\ell$ where \mathcal{A} is a finite alphabet of cardinality κ , and we consider the mutation scheme (1.3). We start by computing M^n . We use the notation and the technique of Section 2.2, and we generalize the formula (2.9) as follows:

$$\begin{aligned} M^n(u, v) &= \mathbb{P}(X_n = v \mid X_0 = u) = \mathbb{P}(X_n(i) = v(i), 1 \leq i \leq \ell \mid X_0 = u) \\ &= \prod_{1 \leq i \leq \ell} \mathbb{P}(X_n(i) = v(i) \mid X_0(i) = u(i)). \end{aligned} \quad (4.11)$$

Now, we have

$$\mathbb{P}(X_n(i) = v(i) \mid X_0(i) = u(i)) = \begin{cases} \frac{1}{\kappa} + \frac{\kappa-1}{\kappa} \rho^n & \text{if } u(i) = v(i), \\ \frac{1}{\kappa} (1 - \rho^n) & \text{if } u(i) \neq v(i), \end{cases} \tag{4.12}$$

where

$$\rho = 1 - \frac{\kappa q}{\kappa - 1}.$$

Coming back to the formula (4.11), we conclude that

$$M^n(u, v) = \left(\frac{1}{\kappa} + \frac{\kappa - 1}{\kappa} \rho^n\right)^{\ell-d(u,v)} \left(\frac{1}{\kappa} (1 - \rho^n)\right)^{d(u,v)}. \tag{4.13}$$

We wish now to compute

$$\sum_{n \geq 1} \frac{M^n}{\lambda^n}.$$

Our first strategy consists of expanding the formula (4.13). We end up with several geometric series, from which we obtain a closed finite formula. So, starting from (4.13) and setting $d = d(u, v)$ to simplify the notation, we have

$$M^n(u, v) = \sum_{i=0}^{\ell-d} \sum_{j=0}^d \binom{\ell-d}{i} \left(\frac{1}{\kappa}\right)^{\ell-d-i} \left(\frac{\kappa-1}{\kappa}\right)^i \rho^{ni} \binom{d}{j} \frac{1}{\kappa^d} (-1)^j \rho^{nj}. \tag{4.14}$$

We can now sum the geometric series to get

$$\sum_{n \geq 1} \frac{M^n}{\lambda^n}(u, v) = \sum_{i=0}^{\ell-d} \sum_{j=0}^d \binom{\ell-d}{i} \left(\frac{1}{\kappa}\right)^{\ell-d-i} \left(\frac{\kappa-1}{\kappa}\right)^i \binom{d}{j} \frac{1}{\kappa^d} \frac{(-1)^j}{\rho^{i+j} - 1}. \tag{4.15}$$

4.3. Probabilistic interpretation

We provide furthermore a probabilistic interpretation of the apparently complex formula (4.15). Let $S_{\ell-d}$ be a binomial random variable with parameters $\ell - d, 1 - 1/\kappa$. We have

$$\sum_{n \geq 1} \frac{M^n}{\lambda^n}(u, v) = \mathbb{E} \left(\frac{1}{\kappa^d} \sum_{j=0}^d \binom{d}{j} \frac{(-1)^j}{\rho^{S_{\ell-d}+j} - 1} \right). \tag{4.16}$$

The above formula is quite nice. Its drawback is that it contains negative terms, so it is not obvious that the global result is non-negative. We present next a third formula, which avoids this problem.

4.4. Third formula

The trick is to expand the factor $(1 - \rho^n)^d$ in the formula (4.13) in a more complicated way, as follows:

$$M^n(u, v) = \sum_{i=0}^{\ell-d} \binom{\ell-d}{i} \left(\frac{1}{\kappa}\right)^{\ell-d-i} \left(\frac{\kappa-1}{\kappa}\right)^i \rho^{ni} \frac{1}{\kappa^d} (1 - \rho)^d \left(\sum_{j=0}^{n-1} \rho^j\right)^d. \tag{4.17}$$

We wish to compute the series $\sum_{n \geq 1} M^n / \lambda^n$. From now on, we deal with the case $d \geq 1$ (for $d = 0$, we use the expression obtained in (4.15), where all the terms are positive). Focusing on the two terms which depend on n in the formula (4.17), we end up with the series

$$\begin{aligned} \sum_{n \geq 1} \frac{\rho^{ni}}{\lambda^n} \left(\sum_{j=0}^{n-1} \rho^j \right)^d &= \sum_{n \geq 1} \frac{\rho^{ni}}{\lambda^n} \sum_{j_1=0}^{n-1} \dots \sum_{j_d=0}^{n-1} \rho^{j_1+\dots+j_d} \\ &= \sum_{j_1, \dots, j_d \geq 0} \sum_{n > \max(j_1, \dots, j_d)} \frac{\rho^{ni}}{\lambda^n} \rho^{j_1+\dots+j_d} \\ &= \sum_{j_1, \dots, j_d \geq 0} \frac{\left(\frac{\rho^i}{\lambda}\right)^{\max(j_1, \dots, j_d)+1}}{1 - \frac{\rho^i}{\lambda}} \rho^{j_1+\dots+j_d}. \end{aligned} \tag{4.18}$$

Notice that this formula is not quite legitimate for $d = 0$, unless we adopt the convention that for $d = 0$ there is only one term in the sum, corresponding to $j_1 = \dots = j_d = 0$. Therefore it seems safer to perform the above computation only when $d \geq 1$. Let us set $\gamma = \rho^i / \lambda$ and focus on the sum

$$\sum_{j_1, \dots, j_d \geq 0} \gamma^{\max(j_1, \dots, j_d)} \rho^{j_1+\dots+j_d}. \tag{4.19}$$

We shall decompose the sum according to the number r of distinct indices among j_1, \dots, j_d , their different values i_1, \dots, i_r , and the number of times n_1, \dots, n_r each value appears. We get

$$\sum_{r=1}^d \sum_{0 \leq i_1 < \dots < i_r} \sum_{\substack{n_1, \dots, n_r \geq 1 \\ n_1 + \dots + n_r = d}} \sum_{j_1, \dots, j_d \in \mathcal{J}} \gamma^{\max(j_1, \dots, j_d)} \rho^{j_1+\dots+j_d}, \tag{4.20}$$

where the innermost sum extends over the set \mathcal{J} defined by

$$\mathcal{J} = \left\{ j_1, \dots, j_d \geq 0 : \text{for } 1 \leq k \leq d, i_k \text{ appears } n_k \text{ times in } j_1, \dots, j_d \right\}.$$

The number of terms appearing in the last summation corresponds to the number of placements of d balls in r cells with occupancy numbers given by n_1, \dots, n_r (see for instance the classical book of Feller [13, Section II.5]), that is, $d! / (n_1! \dots n_r!)$. For each of these placements, we have

$$\max(j_1, \dots, j_d) = i_r, \quad j_1 + \dots + j_d = n_1 i_1 + \dots + n_r i_r;$$

therefore the formula (4.20) becomes

$$\sum_{r=1}^d \sum_{0 \leq i_1 < \dots < i_r} \sum_{\substack{n_1, \dots, n_r \geq 1 \\ n_1 + \dots + n_r = d}} \frac{d!}{n_1! \dots n_r!} \gamma^{i_r} \rho^{n_1 i_1 + \dots + n_r i_r}. \tag{4.21}$$

We can perform the summation over i_r , which yields

$$\begin{aligned} \sum_{0 \leq i_1 < \dots < i_r} \gamma^{i_r} \rho^{n_1 i_1 + \dots + n_r i_r} &= \\ &= \frac{\gamma \rho^{n_r}}{1 - \gamma \rho^{n_r}} \sum_{0 \leq i_1 < \dots < i_{r-1}} (\gamma \rho^{n_r})^{i_{r-1}} \rho^{n_1 i_1 + \dots + n_{r-1} i_{r-1}}. \end{aligned} \tag{4.22}$$

The number of indices i_1, \dots, i_{r-1} has decreased by 1, and γ has been replaced by $\gamma\rho^{n_r}$. Proceeding in the same way, we perform successively the summations over i_{r-1}, \dots, i_1 until we obtain

$$\sum_{0 \leq i_1 < \dots < i_r} \gamma^{i_r} \rho^{n_1 i_1 + \dots + n_r i_r} = \frac{\gamma \rho^{n_r}}{1 - \gamma \rho^{n_r}} \frac{\gamma \rho^{n_r + n_{r-1}}}{1 - \gamma \rho^{n_r + n_{r-1}}} \dots \frac{\gamma \rho^{n_r + \dots + n_2}}{1 - \gamma \rho^{n_r + \dots + n_2}} \sum_{0 \leq i_1} (\gamma \rho^{n_r + \dots + n_2})^{i_1} \rho^{n_1 i_1}.$$

Recalling that $n_1 + \dots + n_r = d$, we can compute the last sum

$$\sum_{0 \leq i_1} (\gamma \rho^{n_r + \dots + n_2})^{i_1} \rho^{n_1 i_1} = \frac{1}{1 - \gamma \rho^d}$$

and conclude that

$$\sum_{0 \leq i_1 < \dots < i_r} \gamma^{i_r} \rho^{n_1 i_1 + \dots + n_r i_r} = \frac{\gamma \rho^{n_r}}{1 - \gamma \rho^{n_r}} \frac{\gamma \rho^{n_r + n_{r-1}}}{1 - \gamma \rho^{n_r + n_{r-1}}} \dots \frac{\gamma \rho^{n_r + \dots + n_2}}{1 - \gamma \rho^{n_r + \dots + n_2}} \frac{1}{1 - \gamma \rho^d}. \tag{4.23}$$

Putting together the formulas (4.21) and (4.23), we obtain a finite formula for the sum (4.19):

$$\sum_{j_1, \dots, j_d \geq 0} \gamma^{\max(j_1, \dots, j_d)} \rho^{j_1 + \dots + j_d} = \sum_{r=1}^d \sum_{\substack{n_1, \dots, n_r \geq 1 \\ n_1 + \dots + n_r = d}} \frac{d!}{n_1! \dots n_r!} \frac{1}{\gamma \rho^d} \frac{\rho^{m_r + (r-1)n_{r-1} + \dots + n_1}}{\prod_{k=1}^r \left(\frac{1}{\gamma} - \rho^{n_r + \dots + n_{r-k+1}} \right)}. \tag{4.24}$$

Plugging this formula into (4.18) and (4.17) yields that, for $d = d(u, v) \geq 1$,

$$\sum_{n \geq 1} \frac{M^n}{\lambda^n}(u, v) = \sum_{i=0}^{\ell-d} \binom{\ell-d}{i} \left(\frac{1}{\kappa} \right)^{\ell-i} \left(\frac{\kappa-1}{\kappa} \right)^i \left(\frac{1-\rho}{\rho} \right)^d \frac{1}{1 - \frac{\rho^i}{\lambda}} \times \sum_{r=1}^d \sum_{\substack{n_1, \dots, n_r \geq 1 \\ n_1 + \dots + n_r = d}} \frac{d!}{n_1! \dots n_r!} \frac{\rho^{m_r + (r-1)n_{r-1} + \dots + n_1}}{\prod_{k=1}^r \left(\frac{\lambda}{\rho^i} - \rho^{n_r + \dots + n_{r-k+1}} \right)}. \tag{4.25}$$

We can further rewrite this formula using the expectation with respect to a binomial random variable $S_{\ell-d}$ with parameters $\ell - d$ and $1 - 1/\kappa$. Indeed, we have

$$\sum_{n \geq 1} \frac{M^n}{\lambda^n}(u, v) = \frac{1}{\kappa^d} \left(\frac{1-\rho}{\rho} \right)^d \times \mathbb{E} \left(\frac{1}{1 - \frac{\rho^{S_{\ell-d}}}{\lambda}} \sum_{r=1}^d \sum_{\substack{n_1, \dots, n_r \geq 1 \\ n_1 + \dots + n_r = d}} \frac{d!}{n_1! \dots n_r!} \frac{\rho^{m_r + (r-1)n_{r-1} + \dots + n_1}}{\prod_{k=1}^r \left(\frac{\lambda}{\rho^{S_{\ell-d}}} - \rho^{n_r + \dots + n_{r-k+1}} \right)} \right). \tag{4.26}$$

Each of the previous formulas is of interest on its own and may be useful in certain contexts. For instance, with the formula (4.15), we see directly that the coefficients of the matrix N_λ are rational functions of the parameter λ , and that there exists a unique choice for λ such that the Perron–Frobenius eigenvalue of N_λ becomes equal to 1. The analysis of the error threshold conducted in Section 3.2 rests entirely on the probabilistic representation presented in the formula (4.16) in the specific case where $u = v = w^*$ and $d = 0$. Finally, the third formula (4.26) will be useful for analyzing the asymptotic behavior of the non-diagonal entries of the matrix N_λ for the landscape with finitely many peaks, in Section 5.

4.5. Proof of Theorem 2

Let $G : E \rightarrow \mathbb{R}$ be an additive functional, that is, a function given by

$$\forall u \in E \quad G(u) = \sum_{i=1}^{\ell} g(u(i)),$$

where g is a function defined on \mathcal{A} with values in \mathbb{R} . We consider an arbitrary fitness function satisfying Hypothesis 1, and we denote by $(x(u), u \in E)$ the solution of the quasispecies equation associated with f . Our goal here is to relate the mean value \bar{G} of G in the quasispecies $(x(u), u \in E)$, defined by

$$\bar{G} = \sum_{u \in E} G(u)x(u),$$

to the mean fitness λ or \bar{f} , defined by

$$\bar{f} = \sum_{u \in E} f(u)x(u).$$

Our starting point is the formula (4.6), which yields

$$\begin{aligned} \bar{G} &= \sum_{u \in E} G(u) \sum_{v \in E} x(v)(f(v) - 1) \sum_{n \geq 1} \frac{1}{\lambda^n} M^n(v, u) \\ &= \sum_{v \in E} x(v)(f(v) - 1) \sum_{n \geq 1} \frac{1}{\lambda^n} \sum_{u \in E} M^n(v, u) G(u). \end{aligned} \quad (4.27)$$

We compute next the innermost sum. By the definition of G ,

$$\sum_{u \in E} M^n(v, u) G(u) = \sum_{i=1}^{\ell} \sum_{u \in E} M^n(v, u) g(u(i)). \quad (4.28)$$

Moreover, the formulas (4.11) and (4.12) yield that, for $1 \leq i \leq \ell$,

$$\begin{aligned} \sum_{u \in E} M^n(v, u) g(u(i)) &= \left(\frac{1}{\kappa} + \frac{\kappa - 1}{\kappa} \rho^n \right) g(v(i)) + \left(\frac{1}{\kappa} (1 - \rho^n) \right) \sum_{a \in \mathcal{A} \setminus \{v(i)\}} g(a) \\ &= \rho^n g(v(i)) + (1 - \rho^n) \bar{g}, \end{aligned} \quad (4.29)$$

where \bar{g} is the mean value of g over the alphabet \mathcal{A} , i.e.,

$$\bar{g} = \frac{1}{\kappa} \sum_{a \in \mathcal{A}} g(a). \quad (4.30)$$

Plugging the formula (4.29) into the formula (4.28), we obtain

$$\sum_{u \in E} M^n(v, u) G(u) = \rho^n G(v) + \ell(1 - \rho^n) \bar{g}. \tag{4.31}$$

Inserting this last formula into (4.27), we obtain

$$\begin{aligned} \bar{G} &= \sum_{v \in E} x(v)(f(v) - 1) \sum_{n \geq 1} \frac{1}{\lambda^n} \left(\rho^n G(v) + \ell(1 - \rho^n) \bar{g} \right) \\ &= \sum_{v \in E} x(v)(f(v) - 1) \left(\frac{\rho}{\lambda - \rho} G(v) + \frac{\ell \bar{g}}{\lambda - 1} - \frac{\ell \bar{g} \rho}{\lambda - \rho} \right) \\ &= \frac{\rho}{\lambda - \rho} (\bar{f} \bar{G} - \bar{G}) + (\bar{f} - 1) \ell \bar{g} \frac{\lambda(1 - \rho)}{(\lambda - 1)(\lambda - \rho)}, \end{aligned} \tag{4.32}$$

where we have introduced the notation

$$\bar{f} \bar{G} = \sum_{u \in E} f(u) G(u) x(u). \tag{4.33}$$

Recalling that $\bar{f} = \lambda$, we have proved the curious formula stated in Theorem 2.

5. Finitely many peaks

As we did for the sharp-peak landscape, we shall give two proofs of Theorem 3: a soft one, whose starting point is the initial quasispecies equation, and a technical one, based on the strategy explained at the end of Section 4. Again, the soft proof might look easier, but the technical proof is more informative and could potentially yield more accurate estimates. We define first the set W^* of the wild types as

$$W^* = \{ w_i : 1 \leq i \leq k \}.$$

Throughout the proofs, we write simply $\lambda(\ell, q)$ or even λ instead of $\lambda_{FP}(\ell, q)$.

5.1. Soft proof

We start as in the case of the sharp-peak landscape. In the long-chain regime, we have that $\sigma(1 - q)^\ell \rightarrow \sigma e^{-a}$. Using the lower bound (4.10) and the fact that $\lambda \geq 1$, we see that

$$\liminf \lambda \geq \max(1, \sigma e^{-a}). \tag{5.1}$$

Suppose that λ does not converge towards $\max(1, \sigma e^{-a})$. In this case, there exist $\varepsilon > 0$ and two sequences of parameters $(\ell_n)_{n \geq 0}, (q_n)_{n \geq 0}$ such that

$$\begin{aligned} \ell_n &\rightarrow +\infty, & q_n &\rightarrow 0, & \ell_n q_n &\rightarrow a, \\ \forall n \geq 0 & \quad \lambda(\ell_n, q_n) &\geq \max(1, \sigma e^{-a}) + \varepsilon. \end{aligned} \tag{5.2}$$

For $1 \leq i \leq k$, we bound from above the quasispecies equation (1.1) associated with w_i^* , as follows:

$$\lambda x(w_i^*) \leq \sum_{v \in W^*} \sigma x(v) M(v, w_i^*) + \sum_{v \in E \setminus W^*} x(v) M(v, w_i^*). \quad (5.3)$$

We sum Equation (5.3) over $i \in \{1, \dots, k\}$. Setting

$$x(W^*) = \sum_{1 \leq i \leq k} x(w_i^*),$$

$$\forall v \in E \quad M(v, W^*) = \sum_{1 \leq i \leq k} M(v, w_i^*),$$

we obtain

$$\lambda x(W^*) \leq \sum_{v \in W^*} \sigma x(v) M(v, W^*) + \sum_{v \in E \setminus W^*} x(v) M(v, W^*). \quad (5.4)$$

Next, using the fact that $M(v, w_i^*) \leq q$ whenever $v \neq w_i^*$, we obtain that

$$\forall v \in W^* \quad M(v, W^*) \leq M(v, v) + kq = (1 - q)^\ell + kq,$$

$$\forall v \in E \setminus W^* \quad M(v, W^*) \leq kq.$$

Inserting these inequalities into (5.4), we obtain that

$$\begin{aligned} \lambda x(W^*) &\leq \sigma x(W^*)((1 - q)^\ell + kq) + (1 - x(W^*))kq \\ &\leq \sigma x(W^*)(1 - q)^\ell + \sigma kq. \end{aligned} \quad (5.5)$$

In addition, we have that

$$\lambda \leq \sigma x(W^*) + 1 - x(W^*) \leq \sigma x(W^*) + 1. \quad (5.6)$$

Equation (5.6) and the condition (5.2) yield that

$$\frac{1}{x(W^*)} \leq \frac{\sigma}{\lambda - 1} \leq \frac{\sigma}{\varepsilon}. \quad (5.7)$$

Using (5.5) and (5.7) together, we obtain that

$$\forall n \geq 0 \quad \lambda(\ell_n, q_n) \leq \sigma(1 - q_n)^{\ell_n} + \frac{\sigma}{\varepsilon} \sigma k q_n.$$

Sending n to ∞ , we see that this inequality is not coherent with the inequality (5.2). Therefore it must be the case that λ converges towards $\max(1, \sigma e^{-a})$.

5.2. Technical proof

For this proof, we shall implement the strategy explained at the end of Section 4. So we consider the matrix $(N_\lambda(i, j))_{1 \leq i, j \leq k}$ defined by

$$\forall i, j \in \{1, \dots, k\} \quad N_\lambda(i, j) = \left((F - I) \sum_{n \geq 1} \frac{M^n}{\lambda^n} \right) (w_i^*, w_j^*), \quad (5.8)$$

where the matrix F was introduced in the formula (4.4). The advantage is that the size of the matrix N_λ is $k \times k$ and this size does not vary with ℓ and q . The diagonal elements of the matrix are given by

$$N_\lambda(i, i) = (\sigma_i - 1) \sum_{n \geq 1} \frac{1}{\lambda^n} M^n(w_i^*, w_i^*), \quad 1 \leq i \leq k.$$

In fact, the above series does not depend on i . Using the formula (4.16) with $d = 0$, we have

$$\forall i \in \{1, \dots, k\} \quad N_\lambda(i, i) = (\sigma_i - 1) \mathbb{E} \left(\frac{1}{\frac{\lambda}{\rho^{S_\ell}} - 1} \right), \quad (5.9)$$

where S_ℓ is a binomial random variable with parameters $\ell, 1 - 1/\kappa$. The parameter λ is adjusted so that the Perron–Frobenius eigenvalue of the matrix N_λ is equal to 1. This Perron–Frobenius eigenvalue is always larger than or equal to any diagonal element; thus we must have $N_\lambda(i, i) \leq 1$. Proceeding exactly as in the technical proof of Theorem 1, we obtain with the help of Jensen’s inequality that

$$\forall i \in \{1, \dots, k\} \quad \lambda \geq \sigma_i \rho^{(\kappa-1)\ell/\kappa}. \quad (5.10)$$

Taking the maximum over $i \in \{1, \dots, k\}$, we get

$$\lambda \geq \sigma \rho^{(\kappa-1)\ell/\kappa}. \quad (5.11)$$

Next we seek an upper bound on the value of λ . The Perron–Frobenius eigenvalue of the matrix N_λ is equal to 1, and all its entries are non-negative; thus

$$1 \leq \max_{1 \leq i \leq k} \sum_{1 \leq j \leq k} N_\lambda(i, j).$$

In particular, there exists an index $i \in \{1, \dots, k\}$ such that

$$1 \leq \sum_{1 \leq j \leq k} N_\lambda(i, j). \quad (5.12)$$

Let us study the non-diagonal elements of the matrix N_λ . For $i \neq j$, we have

$$N_\lambda(i, j) = (\sigma_i - 1) \sum_{n \geq 1} \frac{1}{\lambda^n} M^n(w_i^*, w_j^*).$$

In fact, the series depends only on the Hamming distance between w_i^* and w_j^* , as we can see from the formula (4.13). The same formula shows also that $M^n(w_i^*, w_j^*)$ is a non-increasing function of $d(w_i^*, w_j^*)$. Therefore, using the formula (4.26) with $d = 1$, we have the inequality

$$N_\lambda(i, j) \leq (\sigma_i - 1) \frac{\lambda(1 - \rho)}{\kappa} \mathbb{E} \left(\frac{1}{\rho^{S_{\ell-1}} \left(\frac{\lambda}{\rho^{S_{\ell-1}}} - 1 \right) \left(\frac{\lambda}{\rho^{S_{\ell-1}}} - \rho \right)} \right), \quad (5.13)$$

where $S_{\ell-1}$ is a binomial random variable with parameters $\ell - 1, 1 - 1/\kappa$. Moreover we know that $1 \leq \lambda \leq \sigma$, so

$$N_\lambda(i, j) \leq (\sigma_i - 1) \sigma \frac{(1 - \rho)}{\kappa} \mathbb{E} \left(\frac{1}{\left(\frac{\lambda}{\rho^{S_{\ell-1}}} - 1 \right) \left(\lambda - \rho^{S_{\ell-1}+1} \right)} \right). \quad (5.14)$$

Plugging (5.9) and (5.14) into the inequality (5.12), and recalling that $\sigma \geq \sigma_i$, we get

$$\frac{1}{\sigma - 1} \leq \mathbb{E}\left(\frac{1}{\frac{\lambda}{\rho^{S_\ell}} - 1}\right) + k\sigma \frac{(1 - \rho)}{\kappa} \mathbb{E}\left(\frac{1}{\left(\frac{\lambda}{\rho^{S_{\ell-1}}} - 1\right)(\lambda - \rho^{S_{\ell-1}+1})}\right). \quad (5.15)$$

We proceed in a way similar to what we did for the sharp-peak landscape. Namely, we pick a positive number ε and split the expectations according to the values of S_ℓ and $S_{\ell-1}$, as follows:

$$\begin{aligned} \frac{1}{\sigma - 1} &\leq \frac{1}{\lambda - 1} \mathbb{P}\left(\left|S_\ell - \frac{\kappa - 1}{\kappa} \ell\right| > \varepsilon \ell\right) + \frac{1}{\frac{\lambda}{\rho^{(\kappa-1)\ell/\kappa - \varepsilon \ell}} - 1} \\ &\quad + \frac{k\sigma(1 - \rho)}{\kappa(\lambda - 1)(\lambda - \rho)} \mathbb{P}\left(\left|S_{\ell-1} - \frac{\kappa - 1}{\kappa}(\ell - 1)\right| > \varepsilon(\ell - 1)\right) \\ &\quad + \frac{k\sigma(1 - \rho)}{\kappa} \frac{1}{\left(\frac{\lambda}{\rho^{(\kappa-1)(\ell-1)/\kappa - \varepsilon \ell}} - 1\right)(\lambda - \rho^{(\kappa-1)(\ell-1)/\kappa - \varepsilon \ell})}. \end{aligned} \quad (5.16)$$

Using the estimate given by Chebyshev's inequality (3.3), we obtain

$$\begin{aligned} \frac{1}{\sigma - 1} &\leq \frac{1}{\lambda - 1} \frac{\kappa - 1}{\kappa^2 \varepsilon^2 \ell} + \frac{1}{\frac{\lambda}{\rho^{(\kappa-1)\ell/\kappa - \varepsilon \ell}} - 1} + \frac{k\sigma(1 - \rho)}{\kappa(\lambda - 1)(\lambda - \rho)} \frac{\kappa - 1}{\kappa^2 \varepsilon^2 (\ell - 1)} \\ &\quad + \frac{k\sigma(1 - \rho)}{\kappa} \frac{1}{\left(\frac{\lambda}{\rho^{(\kappa-1)(\ell-1)/\kappa - \varepsilon \ell}} - 1\right)(\lambda - \rho^{(\kappa-1)(\ell-1)/\kappa - \varepsilon \ell})}. \end{aligned} \quad (5.17)$$

Finally, regrouping the first and the third term together, as well as the second and the fourth term, and using that $1 - \rho \leq \lambda - \rho$ and $\rho \leq 1$, we get

$$\begin{aligned} \frac{1}{\sigma - 1} &\leq \frac{\kappa + k\sigma}{\kappa(\lambda - 1)} \frac{\kappa - 1}{\kappa^2 \varepsilon^2 (\ell - 1)} \\ &\quad + \frac{1}{\frac{\lambda}{\rho^{(\kappa-1)(\ell-1)/\kappa - \varepsilon \ell}} - 1} \left(1 + \frac{\frac{k\sigma(1 - \rho)}{\kappa}}{\left(\lambda - \rho^{(\kappa-1)(\ell-1)/\kappa - \varepsilon \ell}\right)}\right). \end{aligned} \quad (5.18)$$

By making an appropriate choice for ε , we shall now deduce an upper bound on λ from this inequality.

Lemma 2. *There exists a positive constant c such that, in the long-chain regime, the mean fitness λ satisfies*

$$\lambda \leq \max\left(1 + \frac{c}{\ell^{1/4}}, \sigma\left(1 - \frac{\kappa q}{\kappa - 1}\right)^{(\kappa-1)(\ell-1)/\kappa - \ell^{2/3}} \left(1 + \frac{1}{\ell^{1/13}}\right)\right). \quad (5.19)$$

Before proving Lemma 2, we explain the end of the technical proof. The lower bound (5.11) on λ and the upper bound (5.19) on λ yield the desired conclusion, which terminates the technical proof. We have also obtained estimates on the speed of convergence of λ , which could be improved with some additional work.

Proof of lemma 2. Suppose that λ is larger than the upper bound in (5.19). We check that this is not compatible with the inequality (5.18) with the choice $\varepsilon = \ell^{-1/3}$. Indeed, we have on

the one hand

$$\lambda > 1 + \frac{c}{\ell^{1/4}} \implies \frac{\kappa + k\sigma}{\kappa(\lambda - 1)} \frac{\kappa - 1}{\kappa^2 \ell^{-2/3} (\ell - 1)} = O\left(\frac{1}{\ell^{1/12}}\right).$$

On the other hand,

$$\lambda > \sigma \left(1 - \frac{\kappa q}{\kappa - 1}\right)^{(\kappa-1)(\ell-1)/\kappa - \ell^{2/3}} \left(1 + \frac{1}{\ell^{1/13}}\right) \implies$$

$$\exists c' > 0 \quad \frac{1}{\frac{\lambda}{\rho^{(\kappa-1)(\ell-1)/\kappa - \ell^{2/3}}} - 1} \leq \frac{1}{\sigma - 1} - \frac{c'}{\ell^{1/13}} + o\left(\frac{1}{\ell^{1/13}}\right).$$

We have also that

$$\frac{\frac{k\sigma(1-\rho)}{\kappa}}{\left(\lambda - \rho^{(\kappa-1)(\ell-1)/\kappa - \ell^{2/3}}\right)} = O\left(\frac{1}{\ell}\right)$$

(notice that we use here the fact that a is finite). Combining the previous inequalities, we find that the right-hand side of the inequality (5.18) has the following expansion:

$$O\left(\frac{1}{\ell^{1/12}}\right) + \frac{1}{\sigma - 1} - \frac{c'}{\ell^{1/13}} + o\left(\frac{1}{\ell^{1/13}}\right) + O\left(\frac{1}{\ell}\right).$$

This quantity becomes strictly less than $1/(\sigma - 1)$ for ℓ large enough, and this is not compatible with the inequality (5.18). □

6. Plateau

This section is devoted to the proof of Theorems 4 and 5. We first show that the value $\lambda(a, \sigma)$ is well-defined. Then we perform a lumping procedure and study the asymptotics of the reduced equations. This yields the convergence of λ towards $\lambda(a, \sigma)$. The other claims of the theorem are proved by similar arguments.

6.1. Definition of $\lambda(a, \sigma)$

We start by proving that Equation (1.13) admits a unique solution. Let us define the function ϕ_n by

$$\phi_n(a) = e^{-an} \sum_{k \geq 0} \frac{(an)^{2k}}{2^{2k} (k!)^2}. \tag{6.1}$$

For any $a > 0$, we have

$$\phi_n(a) = e^{-an} \sum_{k \geq 0} \left(\frac{1}{k!} \left(\frac{an}{2}\right)^k\right)^2 \leq e^{-an} \left(\sum_{k \geq 0} \frac{1}{k!} \left(\frac{an}{2}\right)^k\right)^2 = 1.$$

It follows that the map

$$\lambda \mapsto \sum_{n \geq 1} \frac{\phi_n(a)}{\lambda^n}$$

is continuous and decreasing on $]1, +\infty[$. Moreover it converges to 0 when λ goes to ∞ and to $+\infty$ when λ goes to 1, because

$$\begin{aligned} \sum_{n \geq 1} \phi_n(a) &\geq \sum_{n \geq 1} e^{-an} \sum_{k \geq 0} \frac{(an)^{2k}}{(2k+1)!} \\ &= \sum_{n \geq 1} \frac{e^{-an}}{an} \sum_{k \geq 0} \frac{(an)^{2k+1}}{(2k+1)!} = \sum_{n \geq 1} \frac{e^{-an}}{an} \operatorname{sh}(an) = +\infty. \end{aligned}$$

Thus there exists a unique value $\lambda(a, \sigma)$ satisfying Equation (1.13). We prove next the main claim of the theorem. Throughout the proof, we write simply $\lambda(\ell, q)$ or even λ instead of $\lambda_{\text{PL}}(\ell, q)$.

6.2. Lumping

We are dealing with a fitness function which depends only on the Hamming classes. So we lump together the genotypes according to their Hamming classes and introduce the new variables $y(h)$, $0 \leq h \leq \ell$, given by

$$\forall h \in \{0, \dots, \ell\} \quad y(h) = \sum_{u: H(u)=h} x(u).$$

Using the formula (4.6), we have

$$\begin{aligned} \forall h \in \{0, \dots, \ell\} \quad y(h) &= \sum_{u: H(u)=h} \sum_v x(v)(f(v) - 1) \sum_{n \geq 1} \frac{1}{\lambda^n} M^n(v, u) \\ &= \sum_v x(v)(f(v) - 1) \sum_{n \geq 1} \frac{1}{\lambda^n} \sum_{u: H(u)=h} M^n(v, u). \quad (6.2) \end{aligned}$$

Now, the point is that the innermost sum depends only on the Hamming class of v . Indeed, let us fix $b \in \{0, \dots, \ell\}$ and let v be such that $H(v) = b$. After n mutation steps starting from v , we obtain a genotype X_n whose components are ℓ independent Bernoulli random variables. Among them, exactly b have parameter $(1 + \rho^n)/2$, and $\ell - b$ have parameter $(1 - \rho^n)/2$ (this is a consequence of the computations performed in Section 2.2). Therefore the Hamming class of X_n is distributed as the sum of two independent binomial random variables with respective parameters $(b, (1 + \rho^n)/2)$ and $(\ell - b, (1 - \rho^n)/2)$. We conclude that

$$\begin{aligned} \forall b, c \in \{0, \dots, \ell\}, \quad \forall v \in \{0, 1\}^\ell, \\ H(v) = b \implies \sum_{u: H(u)=c} M^n(v, u) = M_H^n(b, c), \quad (6.3) \end{aligned}$$

where the matrix $M_H^n(b, c)$ is given by

$$M_H^n(b, c) = \mathbb{P}\left(\operatorname{Bin}\left(b, \frac{1 + \rho^n}{2}\right) + \operatorname{Bin}\left(\ell - b, \frac{1 - \rho^n}{2}\right) = c\right), \quad (6.4)$$

$\operatorname{Bin}(n, p)$ being a generic binomial random variable with parameters (n, p) , and the two binomial random variables appearing above being independent. From the definition (1.12), we see

that the fitness f_{PL} depends only on the Hamming class, i.e., there exists a function f such that $f_{\text{PL}}(v) = f(H(v))$ for any v . With a slight abuse of notation, we denote this function also by f_{PL} . Rearranging the sum in the formula (6.2) according to the Hamming class of v and using (6.3), we get

$$\begin{aligned} \forall h \in \{0, \dots, \ell\} \quad y(h) &= \sum_{0 \leq k \leq \ell} \sum_{v: H(v)=k} x(v)(f_{\text{PL}}(v) - 1) \sum_{n \geq 1} \frac{1}{\lambda^n} M_H^n(k, h) \\ &= \sum_{0 \leq k \leq \ell} y(k)(f_{\text{PL}}(k) - 1) \sum_{n \geq 1} \frac{1}{\lambda^n} M_H^n(k, h). \end{aligned} \tag{6.5}$$

Equation (6.5) is in fact valid as long as the fitness function depends only on the Hamming classes. In the case of the plateau landscape, the equation for $h = \ell/2$ yields

$$\frac{1}{\sigma - 1} = \sum_{n \geq 1} \frac{1}{\lambda^n} M_H^n\left(\frac{\ell}{2}, \frac{\ell}{2}\right). \tag{6.6}$$

This is the counterpart of Equation (2.7) for the sharp-peak landscape.

6.3. Asymptotics of $M_H^n(\ell/2, \ell/2)$

We shall rely on Equation (6.6) to analyze the asymptotic behavior of λ in the long-chain regime. Let us begin with the term $M_H^n(\ell/2, \ell/2)$.

Lemma 3. *In the long-chain regime, we have the convergence*

$$\forall n \geq 1 \quad \lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} M_H^n\left(\frac{\ell}{2}, \frac{\ell}{2}\right) = \phi_n(a), \tag{6.7}$$

where the sequence of functions ϕ_n is defined in (6.1).

Proof. From the formula (6.4) with $b = c = \ell/2$, we have

$$\begin{aligned} M_H^n\left(\frac{\ell}{2}, \frac{\ell}{2}\right) &= \mathbb{P}\left(\text{Bin}\left(\frac{\ell}{2}, \frac{1 + \rho^n}{2}\right) + \text{Bin}\left(\frac{\ell}{2}, \frac{1 - \rho^n}{2}\right) = \frac{\ell}{2}\right) \\ &= \mathbb{P}\left(\text{Bin}\left(\frac{\ell}{2}, \frac{1 - \rho^n}{2}\right) = \text{Bin}\left(\frac{\ell}{2}, \frac{1 - \rho^n}{2}\right)\right), \end{aligned}$$

with the understanding that the two binomial random variables appearing above are independent. Moreover, for $a > 0$ we have

$$\frac{\ell}{2} \frac{1 - \rho^n}{2} = \frac{\ell}{2} \frac{1 - (1 - 2q)^n}{2} \sim \frac{\ell}{2} nq \sim \frac{na}{2};$$

hence we are in the regime where these binomial laws converge towards the Poisson distribution $\mathcal{P}(na/2)$ with parameter $na/2$. If $a = 0$, the limit distribution is a Dirac mass at 0. Let us fix $K \geq 1$. We have

$$M_H^n\left(\frac{\ell}{2}, \frac{\ell}{2}\right) \geq \sum_{0 \leq k \leq K} \mathbb{P}\left(\text{Bin}\left(\frac{\ell}{2}, \frac{1 - \rho^n}{2}\right) = k\right)^2,$$

whence, passing to the limit,

$$\liminf_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} M_H^n\left(\frac{\ell}{2}, \frac{\ell}{2}\right) \geq e^{-an} \sum_{0 \leq k \leq K} \frac{(an)^{2k}}{2^{2k}(k!)^2}.$$

Sending K to ∞ , we obtain that

$$\liminf_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} M_H^n\left(\frac{\ell}{2}, \frac{\ell}{2}\right) \geq e^{-an} \sum_{k \geq 0} \frac{(an)^{2k}}{2^{2k}(k!)^2} = \phi_n(a). \tag{6.8}$$

Let us look for the reverse inequality. We again fix $K \geq 1$ and write

$$M_H^n\left(\frac{\ell}{2}, \frac{\ell}{2}\right) \leq \sum_{0 \leq k \leq K} \mathbb{P}\left(\text{Bin}\left(\frac{\ell}{2}, \frac{1-\rho^n}{2}\right) = k\right)^2 + \mathbb{P}\left(\text{Bin}\left(\frac{\ell}{2}, \frac{1-\rho^n}{2}\right) > K\right).$$

We bound the last term with the help of Markov’s inequality:

$$\mathbb{P}\left(\text{Bin}\left(\frac{\ell}{2}, \frac{1-\rho^n}{2}\right) > K\right) \leq \frac{1}{K} \frac{\ell}{2} \frac{1-\rho^n}{2} \leq \frac{an}{K},$$

where the last inequality holds asymptotically. From the previous two inequalities, we conclude that

$$\limsup_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} M_H^n\left(\frac{\ell}{2}, \frac{\ell}{2}\right) \leq e^{-an} \sum_{0 \leq k \leq K} \frac{(an)^{2k}}{2^{2k}(k!)^2} + \frac{an}{K}.$$

Finally, we send K to ∞ to complete the proof of the lemma. □

6.4. Convergence of λ

Let λ^* be an accumulation point of λ . By definition, there exist two sequences of parameters $(\ell_m)_{m \geq 0}, (q_m)_{m \geq 0}$ such that

$$\ell_m \rightarrow +\infty, \quad q_m \rightarrow 0, \quad \ell_m q_m \rightarrow a, \quad \lambda(\ell_m, q_m) \rightarrow \lambda^*.$$

By Lemma 3 and Fatou’s lemma, we have

$$\begin{aligned} \sum_{n \geq 1} \frac{\phi_n(a)}{(\lambda^*)^n} &= \sum_{n \geq 1} \lim_{m \rightarrow \infty} \frac{1}{\lambda(\ell_m, q_m)^n} M_H^n\left(\frac{\ell_m}{2}, \frac{\ell_m}{2}\right) \\ &\leq \lim_{m \rightarrow \infty} \sum_{n \geq 1} \frac{1}{\lambda(\ell_m, q_m)^n} M_H^n\left(\frac{\ell_m}{2}, \frac{\ell_m}{2}\right) = \frac{1}{\sigma - 1}, \end{aligned} \tag{6.9}$$

where the last equality comes from (6.6). This inequality and the very definition of $\lambda(a, \sigma)$ (see Equation (1.13)) imply that $\lambda^* \geq \lambda(a, \sigma) > 1$. In particular, there exist $\varepsilon > 0$ and $M \geq 1$ such that

$$\forall m \geq M \quad \lambda(\ell_m, q_m) \geq 1 + \varepsilon. \tag{6.10}$$

This last inequality ensures that the series

$$\sum_{n \geq 1} \frac{1}{\lambda(\ell_m, q_m)^n} M_H^n\left(\frac{\ell_m}{2}, \frac{\ell_m}{2}\right)$$

converges uniformly with respect to m , so that we can interchange the limits and the infinite sum to get

$$\begin{aligned} \frac{1}{\sigma - 1} &= \lim_{m \rightarrow \infty} \sum_{n \geq 1} \frac{1}{\lambda(\ell_m, q_m)^n} M_H^n\left(\frac{\ell_m}{2}, \frac{\ell_m}{2}\right) \\ &= \sum_{n \geq 1} \lim_{m \rightarrow \infty} \frac{1}{\lambda(\ell_m, q_m)^n} M_H^n\left(\frac{\ell_m}{2}, \frac{\ell_m}{2}\right) = \sum_{n \geq 1} \frac{\phi_n(a)}{(\lambda^*)^n}. \end{aligned} \tag{6.11}$$

Thus we see that λ^* has to be the solution $\lambda(a, \sigma)$ of Equation (1.13). In conclusion, the only possible accumulation point for λ in the long-chain regime is $\lambda(a, \sigma)$. Therefore λ converges towards $\lambda(a, \sigma)$ as stated, and this completes the proof of Theorem 4.

6.5. Asymptotics of $M_H^n(\lfloor \alpha \ell \rfloor, \lfloor \alpha \ell \rfloor)$

The next lemma is the missing ingredient needed to complete the proof of Theorem 5.

Lemma 4. *In the long-chain regime, we have the convergence*

$$\forall n \geq 1 \quad \lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} M_H^n(\lfloor \alpha \ell \rfloor, \lfloor \alpha \ell \rfloor) = \phi_{n,\alpha}(a), \tag{6.12}$$

where the sequence of functions $\phi_{n,\alpha}$ is defined by

$$\phi_{n,\alpha}(a) = e^{-an} \sum_{k \geq 0} \frac{(\sqrt{4\alpha(1-\alpha)}an)^{2k}}{2^{2k}(k!)^2}. \tag{6.13}$$

Proof. From the formula (6.4) with $b = c = \lfloor \alpha \ell \rfloor$, we have

$$\begin{aligned} M_H^n(\lfloor \alpha \ell \rfloor, \lfloor \alpha \ell \rfloor) &= \mathbb{P}\left(\text{Bin}\left(\lfloor \alpha \ell \rfloor, \frac{1 + \rho^n}{2}\right) + \text{Bin}\left(\ell - \lfloor \alpha \ell \rfloor, \frac{1 - \rho^n}{2}\right) = \lfloor \alpha \ell \rfloor\right) \\ &= \mathbb{P}\left(\text{Bin}\left(\ell - \lfloor \alpha \ell \rfloor, \frac{1 - \rho^n}{2}\right) = \text{Bin}\left(\lfloor \alpha \ell \rfloor, \frac{1 - \rho^n}{2}\right)\right), \end{aligned}$$

with the understanding that the two binomial random variables appearing above are independent. Moreover, we have

$$\ell \frac{1 - \rho^n}{2} = \ell \frac{1 - (1 - 2q)^n}{2} \sim \ell nq \sim na,$$

so we are in the regime where the first binomial law converges towards the Poisson distribution with parameter $(1 - \alpha)na$ while the second binomial law converges towards the Poisson distribution with parameter αna . We proceed as in the proof of Lemma 3 to conclude that

$$\forall n \geq 1 \quad \lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} M_H^n(\lfloor \alpha \ell \rfloor, \lfloor \alpha \ell \rfloor) = e^{-an} \sum_{k \geq 0} \frac{((1 - \alpha)an)^k}{k!} \frac{(\alpha an)^k}{k!},$$

and we rewrite the right-hand quantity as in the formula (6.13). □

6.6. Completion of the proof of Theorem 5

Let $\phi_\alpha(a)$ be the function defined by

$$\phi_\alpha(a) = \sum_{n \geq 1} \phi_{n,\alpha}(a) = \sum_{n \geq 1} e^{-an} \sum_{k \geq 0} \frac{(\sqrt{4\alpha(1-\alpha)an})^{2k}}{2^{2k}(k!)^2}. \tag{6.14}$$

In order to study the behavior of $\phi_\alpha(a)$ as a goes to 0 or ∞ , we take advantage of the simple inequalities $(2k)! \leq 2^{2k}(k!)^2 \leq (2k+1)!$ to bound $\phi_{n,\alpha}(a)$ from above and from below as follows:

$$e^{-an} \frac{\sinh(\sqrt{4\alpha(1-\alpha)an})}{\sqrt{4\alpha(1-\alpha)an}} \leq \phi_{n,\alpha}(a) \leq e^{-an} \cosh(\sqrt{4\alpha(1-\alpha)an}).$$

From these inequalities, we deduce that, when $\alpha \neq 1/2$, we have

$$\lim_{a \rightarrow 0} \phi_\alpha(a) = +\infty, \quad \lim_{a \rightarrow +\infty} \phi_\alpha(a) = 0. \tag{6.15}$$

Moreover, for any $a_0 > 0$, we have

$$\forall a \geq a_0 \quad \phi_{n,\alpha}(a) \leq \exp(-a_0n(1 - \sqrt{4\alpha(1-\alpha)})).$$

Thus, when $\alpha \neq 1/2$, the series $\sum_{n \geq 1} \phi_{n,\alpha}(a)$ is uniformly convergent over $[a_0, +\infty[$ for any $a_0 > 0$, and therefore the function ϕ_α is continuous on $]0, +\infty[$. We consider the equation

$$\frac{1}{\sigma - 1} = \phi_\alpha(a),$$

and we denote by $a_c^1(\sigma, \alpha)$ (respectively $a_c^2(\sigma, \alpha)$) the smallest (respectively the largest) solution to this equation in $]0, +\infty[$. These two real numbers are well-defined, thanks to the limits (6.15) and the continuity of ϕ_α on $]0, +\infty[$.

We use a strategy similar to that of the proof of Theorem 4. Suppose that $a \geq a_c^2(\sigma, \alpha)$. The equation for $h = \lfloor \alpha \ell \rfloor$ in the system (6.5) yields

$$\frac{1}{\sigma - 1} = \sum_{n \geq 1} \frac{1}{\lambda^n} M_H^n(\lfloor \alpha \ell \rfloor, \lfloor \alpha \ell \rfloor). \tag{6.16}$$

Let λ^* be an accumulation point of λ . If $\lambda^* > 1$, then, proceeding as in the proof of Theorem 4, we pass to the limit along a subsequence in Equation (6.16) to get

$$\frac{1}{\sigma - 1} = \sum_{n \geq 1} \frac{1}{(\lambda^*)^n} \phi_{n,\alpha}(a) < \phi_\alpha(a), \tag{6.17}$$

but this contradicts the fact that $a \geq a_c^2(\sigma, \alpha)$. Suppose that $a < a_c^1(\sigma, \alpha)$. Let λ^* be an accumulation point of λ . Suppose that $\lambda^* = 1$. By Fatou's lemma, we would have

$$\phi_\alpha(a) = \sum_{n \geq 1} \phi_{n,\alpha}(a) \leq \liminf_{\ell \rightarrow \infty, q \rightarrow 0, \ell q \rightarrow a} \sum_{n \geq 1} \frac{1}{\lambda^n} M_H^n(\lfloor \alpha \ell \rfloor, \lfloor \alpha \ell \rfloor) = \frac{1}{\sigma - 1}, \tag{6.18}$$

where the last equality comes from (6.16). Yet the inequality (6.18) stands in contradiction to the fact that $a < a_c^1(\sigma, \alpha)$. Therefore the infimum limit of λ has to be strictly larger than 1.

7. Survival of the flattest

This final section is devoted to the proof of Theorem 6.

7.1. Lower bound on $\lambda_{\text{SP/PL}}(\delta, \sigma, \ell, q)$

The mean fitness $\lambda_{\text{SP/PL}}(\delta, \sigma, \ell, q)$ is also the Perron–Frobenius eigenvalue of the matrix $(f_{\text{SP/PL}}(u)M(u, v), u, v \in E)$. As such, it is a non-decreasing function of the entries of this matrix. Therefore

$$\lambda_{\text{SP/PL}}(\delta, \sigma, \ell, q) \geq \max(\lambda_{\text{SP/PL}}(\delta, 1, \ell, q), \lambda_{\text{SP/PL}}(1, \sigma, \ell, q)).$$

With the help of Theorems 3 and 4, we conclude that

$$\liminf_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} \lambda_{\text{SP/PL}}(\delta, \sigma, \ell, q) \geq \max(\delta e^{-a}, \lambda(a, \sigma)).$$

Let us prove next the reverse inequality.

7.2. Lumping

Let $(x(u), u \in \{0, 1\}^\ell)$ be the solution to the quasispecies equations associated with $f_{\text{SP/PL}}$. We are dealing with a fitness function which depends only on the Hamming classes. So we lump together the genotypes according to their Hamming classes and introduce the new variables $y(h), 0 \leq h \leq \ell$, given by

$$\forall h \in \{0, \dots, \ell\} \quad y(h) = \sum_{u: H(u)=h} x(u).$$

We use the same computations that led us from (6.2) to (6.5) (these computations relied on the fact that the fitness function depended only on the Hamming classes) to conclude that the variables $y(h), 0 \leq h \leq \ell$, satisfy

$$\forall h \in \{0, \dots, \ell\} \quad y(h) = \sum_{0 \leq k \leq \ell} y(k) (f_{\text{SP/PL}}(k) - 1) \sum_{n \geq 1} \frac{1}{\lambda^n} M_H^n(k, h), \quad (7.1)$$

where the matrix $M_H(k, h)$ is defined in (6.4). Again, with a slight abuse of notation, we denote by $f_{\text{SP/PL}}(k)$ the common value of the fitness function $f_{\text{SP/PL}}$ for all the sequences in the Hamming class $H(k)$. In the rest of the proof, we write simply $\lambda(\ell, q)$ or even λ instead of $\lambda_{\text{SP/PL}}(\delta, \sigma, \ell, q)$. The value λ and the variables y also satisfy

$$\lambda = \delta y(0) + \sigma y(\ell/2) + 1 - y(0) - y(\ell/2) \quad (7.2)$$

and

$$\sum_{0 \leq k \leq \ell} y(k) = 1.$$

7.3. The two main equations

Let us rewrite the two equations of the system (7.1) corresponding to $h = 0$ and $h = \ell/2$:

$$\begin{aligned} y(0) &= y(0)(\delta - 1) \sum_{n \geq 1} \frac{1}{\lambda^n} M_H^n(0, 0) + y\left(\frac{\ell}{2}\right)(\sigma - 1) \sum_{n \geq 1} \frac{1}{\lambda^n} M_H^n\left(\frac{\ell}{2}, 0\right), \\ y\left(\frac{\ell}{2}\right) &= y(0)(\delta - 1) \sum_{n \geq 1} \frac{1}{\lambda^n} M_H^n\left(0, \frac{\ell}{2}\right) + y\left(\frac{\ell}{2}\right)(\sigma - 1) \sum_{n \geq 1} \frac{1}{\lambda^n} M_H^n\left(\frac{\ell}{2}, \frac{\ell}{2}\right). \end{aligned} \quad (7.3)$$

Let us examine the possible limits of λ , $y(0)$, and $y(\ell/2)$ along a subsequence in the long-chain regime. Suppose that, along a subsequence, the following convergences take place:

$$\lambda \rightarrow \lambda^*, \quad y(0) \rightarrow y_0^*, \quad y(\ell/2) \rightarrow y_1^*. \quad (7.4)$$

In the remainder of the argument, all the limits are taken along this subsequence. Passing to the limit in (7.2), we get

$$\lambda^* = \delta y_0^* + \sigma y_1^* + 1 - y_0^* - y_1^*. \quad (7.5)$$

Since $\lambda^* \geq \lambda(a, \sigma) > 1$, necessarily $y_0^* + y_1^* > 0$. Moreover, the following convergences take place:

$$M_H^n(0, 0) \rightarrow e^{-na}, \quad M_H^n\left(\frac{\ell}{2}, \frac{\ell}{2}\right) \rightarrow \phi_n(a), \quad (7.6)$$

while

$$M_H^n\left(\frac{\ell}{2}, 0\right) \rightarrow 0, \quad M_H^n\left(0, \frac{\ell}{2}\right) \rightarrow 0. \quad (7.7)$$

The convergences (7.6) have already been proved. The first one is obvious and the second one is the purpose of Lemma 3. To prove the convergences (7.7), we use the expression for M_H computed in (6.4) and get

$$\begin{aligned} M_H^n\left(\frac{\ell}{2}, 0\right) &= \mathbb{P}\left(\text{Bin}\left(\frac{\ell}{2}, \frac{1 + \rho^n}{2}\right) + \text{Bin}\left(\frac{\ell}{2}, \frac{1 - \rho^n}{2}\right) = 0\right) \\ &\leq \mathbb{P}\left(\text{Bin}\left(\frac{\ell}{2}, \frac{1 + \rho^n}{2}\right) = 0\right) \leq \mathbb{P}\left(\text{Bin}\left(\frac{\ell}{2}, \frac{1}{2}\right) = 0\right), \\ M_H^n\left(0, \frac{\ell}{2}\right) &= \mathbb{P}\left(\text{Bin}\left(\ell, \frac{1 - \rho^n}{2}\right) = \frac{\ell}{2}\right) \leq \mathbb{P}\left(\text{Bin}\left(\ell, \frac{1}{2}\right) = \frac{\ell}{2}\right). \end{aligned}$$

It is well known that the right-hand quantities go to 0 as ℓ goes to ∞ .

7.4. Upper bound on $\lambda_{\text{SP/PL}}(\delta, \sigma, \ell, q)$

Proceeding as in the proof of Theorem 4 (see Lemma 3 and thereafter), we take advantage of the fact that $\lambda^* > 1$ to interchange the limit and the four sums appearing in (7.3). So, passing to the limit in (7.3), and using the convergences (7.4), (7.6), (7.7), we get

$$\begin{aligned} y_0^* &= y_0^*(\delta - 1) \sum_{n \geq 1} \frac{e^{-na}}{(\lambda^*)^n} = y_0^* \frac{\delta - 1}{\lambda^* e^a - 1}, \\ y_1^* &= y_1^*(\sigma - 1) \sum_{n \geq 1} \frac{\phi_n(a)}{(\lambda^*)^n}. \end{aligned} \quad (7.8)$$

If $y_0^* > 0$, then the first equation implies that $\lambda^* = \delta e^{-a}$. If $y_1^* > 0$, then the second equation is equivalent to Equation (1.13). Therefore λ^* has to be equal to the unique solution $\lambda(a, \sigma)$ of this equation. In conclusion, there are only two possible limits along a subsequence for λ , namely δe^{-a} and $\lambda(a, \sigma)$; thus

$$\limsup_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} \lambda_{\text{SP/PL}}(\delta, \sigma, \ell, q) \leq \max(\delta e^{-a}, \lambda(a, \sigma)).$$

The final conclusions of the theorem are obtained as a by-product of the previous argument. Indeed, if $\delta e^{-a} \neq \lambda(a, \sigma)$, then it follows from (7.8) that we cannot have simultaneously $y_0^* > 0$ and $y_1^* > 0$. This remark, in conjunction with the identity (7.5), yields the convergences stated at the end of the theorem.

Acknowledgements

The authors are grateful to the referees for their careful reading and their numerous constructive comments: they not only detected several mathematical inaccuracies, but also helped us greatly improve the quality of the paper and its overall presentation.

Funding information

There are no funding bodies to thank in relation to the creation of this article.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

- [1] ANDERSON, J. P., DAIFUKU, R. AND LOEB, L. A. (2004). Viral error catastrophe by mutagenic nucleosides. *Ann. Rev. Microbiol.* **58**, 183–205.
- [2] BRATUS, A. S., NOVOZHILOV, A. S. AND SEMENOV, Y. S. (2014). Linear algebra of the permutation invariant Crow-Kimura model of prebiotic evolution. *Math. Biosci.* **256**, 42–57.
- [3] BRATUS, A. S., NOVOZHILOV, A. S. AND SEMENOV, Y. S. (2014). On the behavior of the leading eigenvalue of Eigen’s evolutionary matrices. *Math. Biosci.* **258**, 134–147.
- [4] BRATUS, A. S., NOVOZHILOV, A. S. AND SEMENOV, Y. S. (2019). Rigorous mathematical analysis of the quasispecies model: from Manfred Eigen to the recent developments. In *Advanced Mathematical Methods in Biosciences and Applications*, Springer, Cham, pp. 27–51.
- [5] CERF, R. AND DALMAU, J. (2016). The distribution of the quasispecies for a Moran model on the sharp peak landscape. *Stoch. Process. Appl.* **126**, 1681–1709.
- [6] CODOÑER, F. M., DARÓS, J.-A., SOLÉ, R. V. AND ELENA, S. F. (2006). The fittest versus the flattest: experimental confirmation of the quasispecies effect with subviral pathogens. *PLOS Pathogens* **2**, 1–7.
- [7] CROTTY, S., CAMERON, C. E. AND ANDINO, R. (2001). RNA virus error catastrophe: direct molecular test by using ribavirin. *Proc. Nat. Acad. Sci. USA* **98**, 6895–6900.
- [8] DOMINGO, E. AND PERALES, C. (2019). Viral quasispecies. *PLOS Genetics* **15**, 1–20.
- [9] DOMINGO, E. AND SCHUSTER, P. (2016). *Quasispecies: From Theory to Experimental Systems*. Springer, Cham.
- [10] EIGEN, M. (1971). Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften* **58**, 465–523.
- [11] EIGEN, M., MCCASKILL, J. AND SCHUSTER, P. (1989). The molecular quasi-species. *Adv. Chem. Phys.* **75**, 149–263.
- [12] FARCI, P. (2011). New insights into the HCV quasispecies and compartmentalization. *Sem. Liver Disease* **31**, 356–374.
- [13] FELLER, W. (1968). *An Introduction to Probability Theory and Its Applications*, Vol. I. John Wiley, New York.

- [14] FRANKLIN, J., LABAR, T. AND ADAMI, C. (2019). Mapping the peaks: fitness landscapes of the fittest and the flattest. *Artificial Life* **25**, 250–262.
- [15] KEMENY, J. G. AND SNELL, J. L. (1976). *Finite Markov Chains*. Springer, New York.
- [16] KINGMAN, J. F. C. (1977). On the properties of bilinear models for the balance between genetic mutation and selection. *Math. Proc. Camb. Phil. Soc.* **81**, 443–453.
- [17] MEYERHANS, A. *et al.* (1989). Temporal fluctuations in HIV quasispecies in vivo are not reflected by sequential HIV isolations. *Cell* **58**, 901–910.
- [18] MORAN, P. A. P. (1976). Global stability of genetic systems governed by mutation and selection. *Math. Proc. Camb. Phil. Soc.* **80**, 331–336.
- [19] SARDANYÉS, J., ELENA, S. F. AND SOLÉ, R. V. (2008). Simple quasispecies models for the survival-of-the-flattest effect: the role of space. *J. Theoret. Biol.* **250**, 560–568.
- [20] TEJERO, H., MARÍN, A. AND MONTERO, F. (2011). The relationship between the error catastrophe, survival of the flattest, and natural selection. *BMC Evolutionary Biol.* **11**, 12 pp.
- [21] WILKE, C. O. *et al.* (2001). Evolution of digital organisms at high mutation rates leads to survival of the flattest. *Nature* **412**, 331–333.