Genetic Algorithms in Time-Dependent Environments

Christopher Ronnewinkel,

temporary address:
Institut für Neuroinformatik,
Ruhr-Universität Bochum,
D-44780 Bochum, Germany

Claus O. Wilke and Thomas Martinetz

Institut für Neuro- und Bioinformatik,
Universität Lübeck,
Ratzeburger Allee 160,
D-23538 Lübeck, Germany

Contact: ronne@neuroinformatik.ruhr-uni-bochum.de

(to be published in the *Proceedings of the 2nd EvoNet Summerschool*,
Natural Computing Series, Springer)

November 2, 1999

Abstract

The influence of time-dependent fitnesses on the infinite population dynamics of simple genetic algorithms (without crossover) is analyzed. Based on general arguments, a schematic phase diagram is constructed that allows one to characterize the asymptotic states in dependence on the mutation rate and the time scale of changes. Furthermore, the notion of *regular* changes is raised for which the population can be shown to converge towards a *generalized* quasispecies. Based on this, error thresholds and an optimal mutation rate are approximately calculated for a generational genetic algorithm with a moving needle-in-the-haystack landscape. The so found phase diagram is fully consistent with our general considerations.

Genetic algorithms (GAs) as special instances of evolutionary algorithms have been established during the last three decades as optimization procedures, but mostly for static problems (see [1] for an overview and [2] for an in-depth presentation of the field). In view of real-world applications, such as routing in data-nets, scheduling, robotics etc., which include essentially dynamic optimization problems, there are two alternative optimization strategies. On the one hand, one can take snapshots of the system and search "offline" for the optimal solutions of the static situation represented by each of these snapshots. In this approach, the algorithm is restarted for every snapshot and solves the new problem from scratch. On the other hand, the optimization algorithm might reevaluate the real, current situation in order to reuse information gained in the past. In this case, the algorithm works "online". As can be argued from the analogies to natural evolution, evolutionary algorithms seem to be promising candidates for "online" optimization [1, 3]. The reevaluation of the situation or environment then introduces a *time-dependency* of the fitness landscape. This time-dependency occurs as external to the algorithm's population and does not emerge from coevolutive interactions. Coevolutive interactions as an alternative source of time dependency in the fitness landscape are not within the scope of this work.

In the last years, many different methods and extensions of standard evolutionary algorithms for the case of time-dependent fitnesses have been analyzed on the basis of experiments (see [3] for a review) but only seldom on the basis of *theoretical* arguments (see [4, 5]). To take a step into the direction of a better theoretical understanding of "online" evolutionary algorithms, we will study the effects of simple time dependencies of the fitness landscape on the dynamics of GAs (without crossover), or more generally saying, of populations under mutation and probabilistic selection. As we will see, it is possible to characterize the asymptotic states of such a system for a particular class of dynamic fitness landscapes that is introduced below. The asymptotic state forms the basis on which it can be decided whether the population is able to adapt to, or track, the changes in the fitness landscape. Our mathematical formalism applies to GAs as well as to biological self-replicating systems, since the analyzed GA model and Eigen's quasispecies model [6, 7, 8] in the molecular evolution theory (see [9] for a recent review) are very similar. Hence, all introduced concepts for GAs are valid and relevant in analogous form for molecular evolutionary systems.

In the following section, we will introduce the model to be analyzed and show the correspondence to the quasispecies model. Then, we will introduce the mathematical framework, based on which we will formally characterize the asymptotic state as fixed point. After presenting the main concepts, we will proceed with the construction of a phase diagram that allows to characterize the order found in the asymptotic state for different parameter settings. Finally, a moving needle-in-the-haystack (NiH) landscape is analyzed and its phase diagram, including the optimal mutation rate, is calculated.

1 Mathematical Framework

In order to study the influence of a time-dependent fitness landscape on the dynamics of a genetic algorithm (GA), we consider GAs to be discrete dynamical systems. A detailed introduction to the resulting dynamical systems model is given by Rowe [10] (in this book). Here, we will only shortly introduce the basic concepts and the notations we use within the present work.

The GA is represented as a generation operator $G_t^{(m)}$ acting on the space Λ_m of all populations of size m for some given encoding of the population members. If we choose the members i to be encoded as bit-strings of length l, this state space is given by

$$\Lambda_m = \{(n_0, \ldots, n_{2^l-1})/m \mid \textstyle\sum_i n_i = m, n_i \in \mathbb{N}_0\},$$

where n_i denotes the number of bit-strings in the population, which are equal to the binary representation of $i \in \{0, \ldots, 2^l - 1\}$.

The generation operator maps the present population onto the next generation,

$$\mathbf{x}(t+1) = G_t^{(m)}[\mathbf{x}(t)].$$

This is achieved by applying a sampling procedure that draws the members of the next generation's population $\mathbf{x}(t+1)$ according to their expected concentrations $\langle \mathbf{x}(t+1) \rangle \in \Lambda_\infty$ which are defined by the mixing [10, 11] and the selection scheme. For an infinite population size, the sampling acts like the identity resulting in

$$G_t^{(\infty)}\mathbf{x}(t) = \mathbf{x}(t+1) = \langle \mathbf{x}(t+1) \rangle.$$

Hence, $G_t := G_t^{(\infty)}$ represents in fact the mixing and selection scheme. For finite population size, $\langle \mathbf{x}(t+1) \rangle \in \Lambda_\infty$ is approximated by using the sampling process to obtain $\mathbf{x}(t+1) \in \Lambda_m$. The deviations thereby possible become larger with decreasing m and distort the finite population dynamics as compared to the infinite population case. This results in fluctuations and epoch formation as shown in [10, 11, 12]. In the following, we will consider the infinite population limit, because it reflects the exact flow of probabilities for a particular fitness landscape. In a second step, the fluctuations and epoch formation introduced by the finiteness of a real population can be studied on the basis of that underlying probability flow.

The generation operator is assumed to decompose into a separate mutation and a separate selection operator, like

$$G_t = M \cdot S(t), \tag{1}$$

where the selection operator $S(t)$ contains the time dependency of the fitness landscape. Crossover is not considered in this work.

Inspired by molecular evolution, and also by common usage, we assume that the mutation acts like flipping each bit with probability μ. If we set the duration of one generation to 1, μ equals to the mutation rate. The mutation operator then takes on the form

$$M = \begin{pmatrix} 1-\mu & \mu \\ \mu & 1-\mu \end{pmatrix}^{\otimes l}, \quad \text{i. e.} \quad M_{ij} = \mu^{d_{\mathrm{H}}(i,j)} (1-\mu)^{l - d_{\mathrm{H}}(i,j)},$$

where $\otimes$ denotes the Kronecker (or canonical tensor) product and $d_{\mathrm{H}}(i,j)$ denotes the Hamming distance of i and j.

To keep the description analytically tractable, we will focus on fitness-proportionate selection,

$$S(t) \cdot \mathbf{x} = F(t) \cdot \mathbf{x} \big/ \langle f(t) \rangle_{\mathbf{x}}, \quad \begin{array}{l} \text{where } F(t) = \operatorname{diag}\big(f_0(t), \ldots, f_{2^l-1}(t)\big) \\ \text{and } \langle f(t) \rangle_{\mathbf{x}} = \sum_i f_i(t) x_i = \| F(t) \cdot \mathbf{x} \|_1. \end{array}$$

This will already provide us with some insight into the general behavior of a GA in time-dependent fitness landscapes.

Since the GA corresponding to Eq. 1 applies mutation to the current population and selects the new population with complete replacement of the current one, it is called a *generational* GA (genGA). In addition to genGAs, *steady-state* GAs (ssGAs) with a two step reproduction process are also in common use: First, a small fraction γ of the current population is chosen to produce $m\gamma$ mutants according to some heuristics. Second, another fraction γ of the current population is chosen to get replaced by those mutants according to some other heuristics (see [14, 15, 16] and references therein). We can include ssGAs into our description in an approximate fashion by simply bypassing a fraction $(1-\gamma)$ of the population into the selection process without mutation, whereas the remaining fraction γ gets mutated before it enters the selection process. The generation operator then reads

$$G_t = \left[(1-\gamma) \mathbb{1} + \gamma M \right] S(t). \tag{2}$$

By varying γ within the interval $]0, 1]$, we can interpolate between steady-state behavior (ssGA) for $\gamma \ll 1$ and generational behavior (genGA) for $\gamma = 1$. Equation 2 is

only an approximation of the true generation operator for ssGAs because the heuristics involved in the choice of the mutated and replaced members are neglected. But in the next section, the heuristics are expected to play a minor role for our general conclusion on an inertia of ssGAs against time-variations.

At this point, we want to review shortly the correspondence of our GA model with the quasispecies model, extensively studied by Eigen and coworkers [6, 7, 8] in the context of molecular evolution theory (see also [13] in this book). The quasispecies model describes a system of self-replicating entities i (e. g. RNA-, DNA-strands) with replication rates f_i and an imperfect copying procedure such that mutations occur. For simplicity reasons, the overall concentration of molecules in the system is held constant by an excess flow $\Phi(t)$. In the above notation, the continuous model reads

$$\dot{\mathbf{x}}(t) = [M \cdot F(t) - \Phi(t)]\,\mathbf{x}(t), \tag{3}$$

where the flux needs to equal the average replication, $\Phi(t) = \langle f(t)\rangle_{\mathbf{x}(t)}$, in order to keep the concentration vector $\mathbf{x}(t)$ normalized. This model might then be discretized via $t \to t/\delta t$, which unveils the similarity to a ssGA:

$$\mathbf{x}(t+1) = \left[(1 - \delta t\,\langle f(t)\rangle_{\mathbf{x}(t)})\mathbb{1} + \delta t\, M \cdot F(t)\right]\mathbf{x}(t) \quad \text{for } \delta t \ll 1. \tag{4}$$

By comparison with Eq. 2, we can easily read off that $\gamma = \delta t\,\langle f(t)\rangle_{\mathbf{x}(t)} =: \gamma_{\mathbf{x}(t)}$. This means a low (resp. high) average fitness leads to a small (resp. large) replacement – a property that is not wanted in the context of optimization problems, which GAs are usually used for, because one does not want to remain in a region of low fitness for a long time. Another difference to ssGAs is the fact that in the continuous Eigen model, selection acts only on the mutated fraction of the population – although this leads only to subtle differences in the dynamics of ssGAs and the Eigen model.

Equation 3 is commonly referred to as 'continuous Eigen model' in the literature, because of the continuous time, and Eq. 4 is simply its discretized form which can be used for numerical calculations. Nonetheless, the notion 'discrete Eigen model' is seldom used for Eq. 4 but it is often used for the genGA,

$$\mathbf{x}(t+1) = [M \cdot S(t)]\,\mathbf{x}(t), \tag{5}$$

in the literature. This stems from the identical asymptotic behavior of Eqs. 4 and 5 for static fitness landscapes. However, there are differences for time-dependent fitness landscapes, as we will see in the following two sections.

2 Regular Changes and Generalized Quasispecies

In the case of a static landscape, the fixed points of the generation operator, which are in fact stationary states of the evolving system (if contained within Λ_m, see [10]), can be found by solving an eigenvalue problem, because of

$$\mathbf{x} = G\mathbf{x} \quad \Longleftrightarrow \quad MF\,\mathbf{x} = \langle f\rangle_{\mathbf{x}}\mathbf{x}\ . \tag{6}$$

Let λ_i and $\mathbf{v}_i$ denote the eigenvalues and eigenvectors of MF with descending order $\lambda_0 \geq \cdots \geq \lambda_{2^l-1}$ and $\|\mathbf{v}_i\|_1 = 1$. For $\mu \neq 0, 1$ the Perron-Frobenius theorem assures

the non-degeneracy of the eigenvector $\mathbf{v}_0$ to the largest eigenvalue and moreover it assures $\mathbf{v}_0 \in \Lambda_\infty$. Often, $\mathbf{v}_0$ is called Perron vector. After a transformation to the basis of the eigenvectors $\{\mathbf{v}_i\}$ it can be straightforwardly shown that $\mathbf{x}(t)$ converges to $\mathbf{v}_0$ for $t \to \infty$. The population represented by $\mathbf{v}_0$ was called the 'quasispecies' by Eigen, because this population does not consist of only a single dominant genotype, or string, but it consists of a particular stable mixture of different genotypes.

Let us now consider time-dependent landscapes. If the time dependency is introduced simply by a single scalar factor, like

$$F(t) = F\,\rho(t) \quad \text{with } \rho(t) \geq 0 \text{ for all } t,$$

it immediately drops out of the selection operator for GAs. For the continuous Eigen model, we note that the eigenvectors of $F(t)$ and F are the same and that $\lambda_i(t) = \lambda_i\,\rho(t)$. Since $\rho(t) \geq 0$, which is necessary to keep the fitness values positive, the order of the eigenvalues remains, such that $MF(t)$ will show the same quasispecies $\mathbf{v}_0$ as MF. Contrasting to that special case, a general, individual time dependency of the string's fitnesses does indeed change the eigenvalues and eigenvectors of $MF(t)$ compared to MF. For an arbitrary time dependency the Perron vector is constantly changing, and therefore, we cannot even define a unique asymptotic state. However, this problem disappears for what we call *regular* changes. After having established a theory for such changes, we can then take into account more and more non-regular ingredients. What do we mean by "*regular* change"? We define it heuristically in the following way: a regular change is a change that happens with fixed duration τ and obeys some deterministic rule that is the same for all change cycles. Let us express the latter more formally and make it more clear what we mean by "same rule of change". Within a change cycle, we allow for an arbitrary time dependency of the fitness, up to the restriction that two different change cycles must be connected by a permutation of the sequence space. Thus, if the time dependency is chosen for one change cycle, e. g. the first change cycle starting at $t = 0$, it is already fixed for all other cycles, apart from the permutations. We will represent permutations π from the permutation group $\mathfrak{S}_{2^l}$ of the sequence space as matrices

$$(P_\pi)_{ij} = \delta_{\pi(i),j} \quad \text{for } i, j \in \{0, \ldots, 2^l - 1\}.$$

The permutations of vectors $\mathbf{x}$ and matrices A are obtained by

$$(P_\pi \mathbf{x})_i = x_{\pi(i)} \quad \text{and} \quad (P_\pi A P_\pi^{\mathrm{T}})_{i,j} = A_{\pi(i),\pi(j)},$$

where P_π^{T} denotes the transpose of P_π with the property $P_\pi^{\mathrm{T}} = P_{\pi^{-1}} = P_\pi^{-1}$.

In reference to the first change cycle, we define the fitness landscape $F(t)$ as being *single-time-dependent*, if and only if for each change cycle $n \in \mathbb{N}_0$ there exists a permutation $\pi_n \in \mathfrak{S}_{2^l}$, such that for all cycle phases $\varphi \in \{0, \ldots, \tau - 1\}$

$$P_n\,F(\varphi + n\tau)\,P_n^{\mathrm{T}} = F(\varphi) \qquad (\text{abbreviatory } P_n := P_{\pi_n}).$$

We will call each permutation P_n a *jump-rule*, or simply *rule*, which connects $F(\varphi+n\tau)$ and $F(\varphi)$. To make predictions about the asymptotic state of the system, we need to relate the generation operators of different change cycles to each other. This is

readily achieved if the permutations P_n *commute* with the mutation operator M. The condition for this being the case is that for all i, j,

$$M_{ij} = M_{\pi_n(i),\pi_n(j)} \quad \text{or equivalently} \quad d_{\mathrm{H}}(i,j) = d_{\mathrm{H}}\big(\pi_n(i), \pi_n(j)\big).$$

Thus, the Hamming distances $d_{\mathrm{H}}(i,j)$ need to be *invariant* under the permutations P_n. Geometrically this means that the fitness landscape gets "translated" or "rotated" by those permutations without changing the neighborhood relations. Then, we find for arbitrary $n \in \mathbb{N}$ and $\varphi \in \{0, \ldots, \tau - 1\}$,

$$G_{\varphi+n\tau} = P_n^{\mathrm{T}} G_\varphi P_n. \tag{7}$$

To study the asymptotic behavior of the system, it is useful to accumulate the time dependency of a change cycle by introducing the τ-generation operators,

$$\Gamma_n := G_{\tau-1+n\tau} \cdots G_{n\tau} \quad \text{for all } n \in \mathbb{N}_0.$$

Because of Eq. 7, all these operators are related to Γ_0 by

$$\Gamma_n = P_n^{\mathrm{T}} \Gamma_0 P_n,$$

This property allows us to write the time evolution of the system in the form

$$\mathbf{x}(\varphi + n\tau) = P_{n-1}^{\mathrm{T}} \Gamma_0 P_{n-1} \ \cdots \ P_1^{\mathrm{T}} \Gamma_0 P_1 \, \Gamma_0 \ \mathbf{x}(\varphi), \tag{8}$$

where $\varphi \in \{0, \ldots, \tau - 1\}$ denotes in the following always the phase within a cycle.

Let us consider the special case of a single rule P being applied at the end of each change cycle, which results in $P_n = (P)^n$, e. g. imagine a fitness peak that moves at a constant "velocity" through the string space. We will see below that for those cases it is possible to identify the asymptotic state with a quasispecies in analogy to static fitness landscapes. Because of that, we can now define the notion of *regularity* of a fitness landscape formally in the following manner:

A time-dependent fitness landscape $F(t)$ is *regular*, if and only if: (i) the fitness landscape is *single-time-dependent*, (ii) there exists some rule $P \in \mathfrak{S}_{2^l}$ which is applied at the end of each cycle such that $P_n = (P)^n$, and (iii) the rule P *commutes* with the mutation operator M.

In this case, we get with $PP^{\mathrm{T}} = \mathbb{1}$ the time evolution

$$\mathbf{x}(\varphi + n\tau) = \big(P^{\mathrm{T}}\big)^n \big(P\Gamma_0\big)^n \, \mathbf{x}(\varphi). \tag{9}$$

To proceed, it is useful to permute the concentrations compatible to the rule of the fitness landscape. By this, concentrations are measured in reference to the fitness landscape structure of the start cycle $n = 0$. We will denote those concentrations by $\mathbf{x}'(t)$ and they are related to the concentrations $\mathbf{x}(t)$ by

$$\begin{aligned} \mathbf{x}'(\varphi + n\tau) &= (P)^n \, \mathbf{x}(\varphi + n\tau) \\ &= (P\Gamma_0)^n \, \mathbf{x}(\varphi) \qquad \text{and} \quad \mathbf{x}'(\varphi) = \mathbf{x}(\varphi). \end{aligned} \tag{10}$$

For example, if there is no time-dependency within the cycles, some x_i' will for all cycles measure the concentration of the highest fitness string, independent of its current

position in string space. Thus, $\mathbf{x}'(t)$ evolves in a fitness landscape with periodic change, which can also be seen from the second line of Eq. 10. In analogy to the static case Eq. 6, the calculation of fixed points of $\mathbf{x}'(t)$ is equivalent to an eigenvalue problem,

$$\mathbf{x}'(t+\tau) = \mathbf{x}'(t) \quad \Longleftrightarrow \quad P\widetilde{\Gamma}_0\,\mathbf{x}'(t) = \|P\widetilde{\Gamma}_0\,\mathbf{x}'(t)\|_1\;\mathbf{x}'(t),$$

where $\widetilde{\Gamma}_0$ is the *unnormalized* τ-generation operator obtained from the accumulation of the *unnormalized* generation operators $\widetilde{G}_\varphi = MF(\varphi)$.

The corresponding periodic quasispecies $\mathbf{v}_0$ can be calculated for all phases φ of the change cycle from the Perron vector $\mathbf{v}_0$ of $P\Gamma_0$ in the following way,

$$\mathbf{x}'(\varphi+n\tau) \xrightarrow{n\to\infty} \mathbf{v}_0(\varphi) = G_{\varphi-1}\cdots G_0\,\mathbf{v}_0 \quad \text{for } \varphi\in\{0,\ldots,\tau-1\}. \tag{11}$$

To find the asymptotic states of the concentrations $\mathbf{x}(t)$, we simply need to invert Eq. 10,

$$\mathbf{x}(\varphi+\nu\tau) = \bigl(P^{\mathrm{T}}\bigr)^{\nu}\mathbf{x}'(\varphi+\nu\tau) \quad \text{for } \nu\in\{0,\ldots,\eta-1\}, \tag{12}$$

where $\eta := \operatorname{ord} P$ is the order of the group element $P\in\mathfrak{S}_{2^l}$.

The essential reason for the existence of asymptotic states for $\mathbf{x}(t)$ lies in the finiteness of the permutation group $\mathfrak{S}_{2^l}$. Because of $P^\eta = \mathbb{1}$, we find directly from Eq. 9 the asymptotic state

$$\mathbf{x}(\varphi+\tilde{n}\eta\,\tau) = (P\Gamma_0)^{\eta\,\tilde{n}}\,\mathbf{x}(t) \xrightarrow{\tilde{n}\to\infty} \mathbf{v}_0(\varphi),$$

where $\mathbf{v}_0(\varphi)$ is the same as in Eq. 11, because $(P\Gamma_0)^\eta$ and $P\Gamma_0$ have the same eigenvectors, in particular the same Perron vector. Moreover, we get

$$\mathbf{x}\bigl(\varphi+(\nu+\tilde{n}\eta)\tau\bigr) \xrightarrow{\tilde{n}\to\infty} \bigl(P^{\mathrm{T}}\bigr)^{\nu}\mathbf{v}_0(\varphi) \quad \text{for } \nu\in\{0,\ldots,\eta-1\}, \tag{13}$$

which is the same result as Eqs. 11 and 12 yield. In the limit of long strings $l\to\infty$, $\operatorname{ord} P$ is not necessarily finite anymore. If $\operatorname{ord} P \xrightarrow{l\to\infty}\infty$, then the asymptotic states Eq. 13 for $\mathbf{x}(t)$ do not exist, but Eq. 11 still holds. Hence, a quasispecies exists even in the limit $l\to\infty$ if measured in reference to the structure of the fitness landscape.

In conclusion, Eqs. 11 and 13 represent the *generalized* quasispecies for the class of *regular* fitness landscapes which includes as special cases static and periodic fitness landscapes. In fact, the simplest case of a *regular* change is a periodic variation of the fitness values $f_i(t) = f_i(t+\tau)$ because *no* permutations are involved ($P=\mathbb{1}$) and hence $\mathbf{x}'(t) = \mathbf{x}(t)$ for all t. The quasispecies was generalized for this case already in [17] and – using a slightly different formalism – in [4]. In Section 4, we will study a more complicated example.

3 Schematic Phase Diagram

To get an intuitive feeling for the typical behavior of ssGAs and genGAs, let us consider some special lines in the plane spanned by the mutation rate μ and the time scale for changes τ, as shown in Fig. 1. The mutation operator represents only for $\mu < 1/2$ a copying procedure with occurring errors, whereas for $\mu > 1/2$ it systematically tends to invert strings, i. e. it resembles an inverter with occurring errors. Since mutation should introduce *weak* modifications to the strings, we will consider only $\mu \le 1/2$.

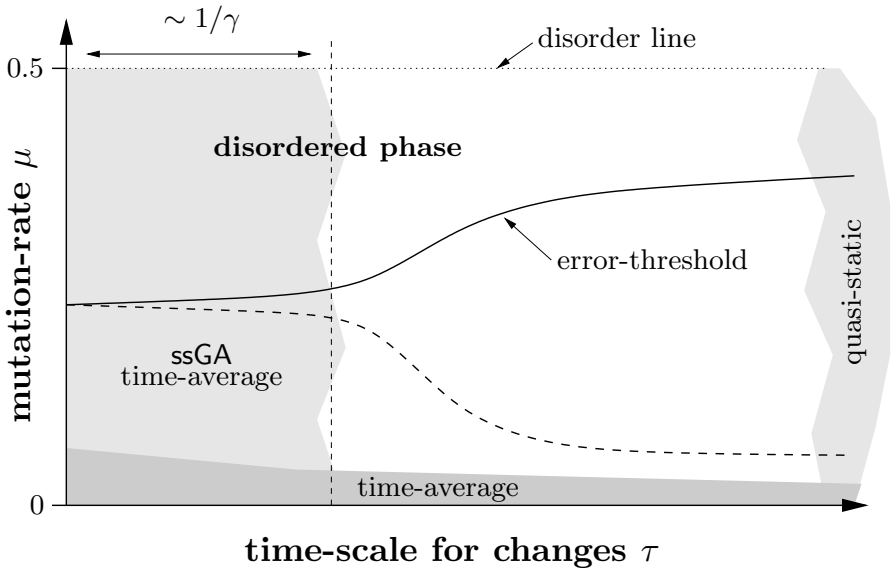


Figure 1: Schematic phase diagram: time-average regions due to low mutation (dark gray) and large inertia (light gray, left), quasi-static region for slow changes (light gray, right).

Disorder line: For $\mu = 1/2$, the Perron vector of $MF(t)$ is always $\mathbf{v}_0^T = (1, \ldots, 1)/2^l$. The population will therefore converge towards the disordered state. Because of the continuity of M in μ, we already enter a disordered phase for $\mu \approx 1/2$.

Time-average region: For $\mu = 0$, the mutation operator is the identity. We find as time evolution simply the product average over the fitness of the evolved time steps:

$$\begin{aligned}\mathbf{x}(t+\tau) &= \left[\prod_{\varphi=t}^{t+\tau-1} S(\varphi)\right]\mathbf{x}(t) \\ &= \tilde{F}(t+\tau,t)\,\mathbf{x}(t)/\|\ldots\|_1 \text{ with } \tilde{F}(t+\tau,t) = \textstyle\prod_{\varphi=t}^{t+\tau-1} F(\varphi).\end{aligned}$$

Since diagonal operators commute, the order in which the $F(\varphi)$ get multiplicated does not make any difference. For the case of a τ-periodic landscape, $\tilde{F} = \tilde{F}(t+\tau,t) = \tilde{F}(\tau,0)$ is *independent* of t. The quasispecies is then a linear superposition of the eigenvectors of the largest eigenvalue of the product averaged fitness landscape $\tilde{F}$ – there might be more then one such eigenvector, since $\tilde{F}$ is diagonal and the Perron-Frobenius theorem does not apply. Because of the continuity of M in μ the dynamics are governed already for $0 < \mu \ll 1$ by the product average $\tilde{F}$. Analogous conclusions apply to those non-periodic landscapes for which by choosing a suitable time scale τ a meaningful average $\tilde{F}(t+\tau,t)$ can be defined.

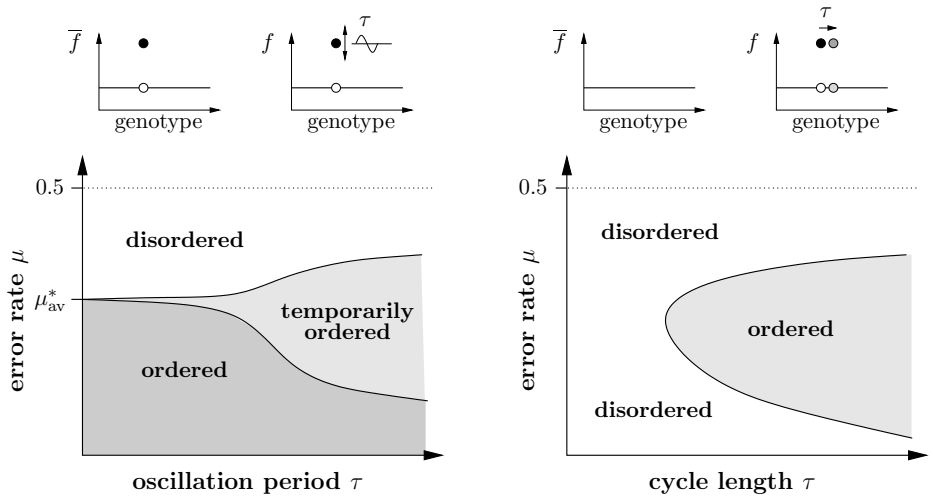


Figure 2: Phase diagrams for *(left)*: needle-in-the-haystack with oscillating height at frequency $\omega = 2\pi/\tau$, *(right)*: needle-in-the-haystack that jumps after τ time steps to a randomly chosen nearest neighbor.

For ssGAs, γ is small and we find to first order in $\tau\gamma$:

$$\mathbf{x}(t+\tau) = (1-\tau\gamma)\tilde{F}(t+\tau,t) + \tau\gamma\left(\frac{1}{\tau}\sum_{\varphi=0}^{\tau-1} S(t+\tau)\cdots\underbrace{M}_{\varphi\text{th factor from left}}\cdots S(t)\right) + \mathcal{O}\big((\tau\gamma)^2\big).$$

If $\tau\gamma \ll 1$ holds, the time evolution is governed by $\tilde{F}(t,t+\tau)$. For changes on a time scale τ, we find time-averaged behavior if $\tau \ll 1/\gamma$. Thus, the width of the time-average region is proportional to $1/\gamma$. A detailed analysis of the effect of the different positions of the mutation operator M within the $\tau\gamma$-term, which is otherwise an arithmetic time-average, has not yet been carried out.

Quasi-static region: If the changes happen on a time scale τ very large compared to the average relaxation time ($\sim 1/\langle\lambda_0 - \lambda_1\rangle$) the quasispecies grows nearly without noticing the changes. Thus, in the quasi-static region all quasispecies that might be expected from the static landscapes $\tilde{F} = F(t)$ will occur at some time during one cycle τ.

Wilke *et al.* raise in [18] the schematic phase diagram of the continuous Eigen model, which exhibits the same time-average phases as that for ssGAs. Their result is in perfect agreement with two recently, explicitly studied time-dependent landscapes. First, Wilke *et al.* studied in [17] a needle-in-the-haystack (NiH) landscape with oscillating, τ-periodic fitness of the needle, i. e.

$$f_0(t) > f_1 = \cdots = f_{2^l-1} = 1 \quad \text{and} \quad f_0(t) = \sigma\exp\{\varepsilon\sin(2\pi\, t/\tau)\}\,.$$

The continuous model was represented for $\delta t \to 0$ as Eq. 4 and the periodic quasispecies Eq. 11 was calculated. Figure 2 *(left)* shows the resulting phase diagram. For small τ,

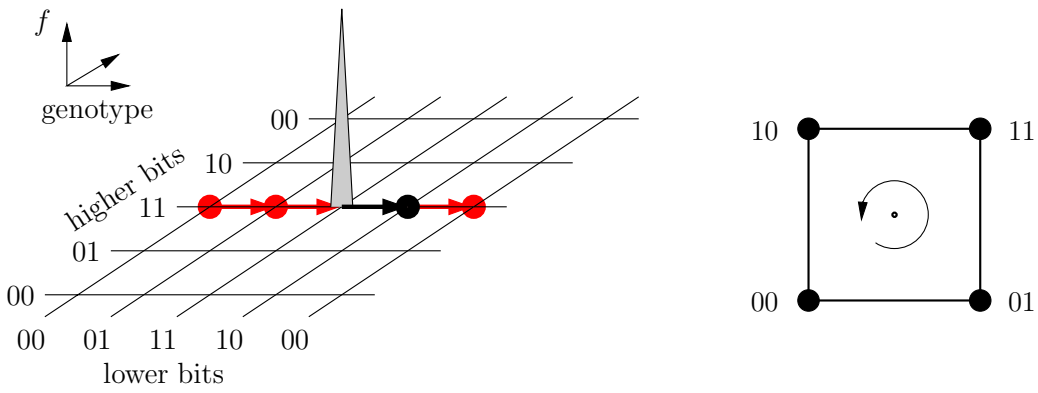


Figure 3: A regularly moving needle-in-the-haystack for string length $l = 4$. In *(left)*, the solid arrow represents the next jump to happen, whereas the gray and solid arrows all together represent the jumps that happen one after the other under the rule P of rotating the two lower bits as shown in *(right)* with rotation angle $\pi/2$ at every jump.

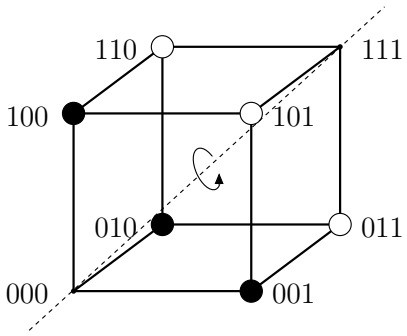
the error threshold is given by the one of the time-averaged landscape, whereas for large τ, the error threshold oscillates between minimum and maximum values corresponding to $\min_t f_0(t)$ and $\max_t f_0(t)$, as expected in the quasi-static regime. Second, Nilsson and Snoad studied in [19] a moving NiH that jumps randomly to one of its nearest neighbor strings every τ time steps. The time-average of this landscape over many jump cycles is a totally flat or neutral landscape, which explains the extension of the disordered phase to small μ and small τ as it is shown in Fig. 2 *(right)*. In the quasi-static region, order is expected because the needle stays long enough at each position for a quasispecies to grow. Hence, we can understand the existence of the observed and calculated phase diagrams in Fig. 2 from simple arguments. In fact, they are special instances of the general schematic phase diagram depicted in Fig. 1.

In the following, we will consider regularly moving NiHs and derive the infinite population behavior of a genGA in such landscapes. This is interesting, since genGAs should be considered to adapt faster to changes compared to ssGAs, as the missing time-average region of genGAs for small τ suggests. To clarify whether a different phase diagram compared to Fig. 2 *(right)* emerges for genGAs with moving NiH, we will calculate the phase diagram including the optimal mutation rate that maximizes a lower bound for the concentration of the needle string in the population.

4 Generational GA and a moving NiH

In this section, we want to analyze quantitatively the asymptotic behavior of a genGA with NiH that moves *regularly* in the sense of Section 2 to one of its l nearest neighbors every τ time steps. At the end, we will also be able to comment on the case of a NiH that jumps *randomly* to one of its nearest neighbors.

A simple example of a NiH that moves regularly to nearest neighbors is shown in Fig. 3 *(left)*. Each jump corresponds to a $\pi/2$-rotation of the four-dimensional hypercube $\{0,1\}^4$ along the 1100 axis, i. e. the lower two bits are rotated as shown in Fig. 3 *(right)*. We will call the set of strings $\{P^n i \mid n \in \mathbb{N}\}$ which is obtained by



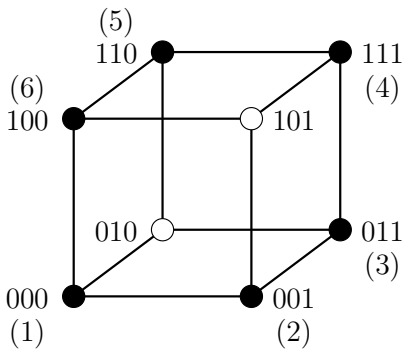
$P_{\ll}$

$000 \to 000$	$111 \to 111$
$001 \to 010$	$011 \to 110$
$010 \to 100$	$110 \to 101$
$100 \to 001$	$101 \to 011$

Figure 4: The equivalence of a $2\pi/3$-rotation along the $1 \cdots 1$ axis and a cyclic 1-bit left-shift, denoted by $P_{\ll}$, for string length $l = 3$.

applying the same rule $P \in \mathfrak{S}_{2^l}$ over and over to some initial string $i \in \{0,1\}^l$, the *orbit of i under P*. The period length 4 of the orbit shown in Fig. 3 *(left)* originates from the rotation angle $\pi/2$ and hence is independent of the string length l. The orbits of such rotations will always be restricted to only four different strings. For reasons that will become clear below, we are looking for *regular* movements of the needle that are *not* restricted to such a small subspace of the string space. Instead, the needle is supposed to move 'straight away' from previous positions in string space. Since a complete classification and analysis of all possible *regular* movements for given string length l and jump distance d is out of the scope of this work, we will simply give an example of a rule $P \in \mathfrak{S}_{2^l}$ that generates such movements: the composition of a cyclic 1-bit left-shift, which we denote by $P_{\ll}$, and an exclusive-or with $0 \cdots 01$, which we denote by $P_{\oplus}$. For string length $l \leq 3$, $P_{\ll}$ corresponds to a $2\pi/l$ rotation along the $1 \cdots 1$ axis as can be seen in Fig. 4. Moreover, the orbit of $0 \cdots 0$ under $P_{\oplus\ll} = P_{\oplus} \circ P_{\ll}$ is shown in Fig. 5 also for $l = 3$. For arbitrary string length l, it is more difficult to visualize the action of $P_{\ll}$ and hence of $P_{\oplus\ll}$. But, it is easily verified that starting from all zeros $0 \cdots 0$, the string with $n \leq l$ ones $0 \cdots 01 \cdots 1$ will be reached after exactly n jumps. Moreover, the orbit of $0 \cdots 0$ under $P_{\oplus\ll}$ has the period length $2l$. In the limit of long strings $l \to \infty$, this periodicity is broken because the needle never (i. e. after ∞ many jumps) returns to all zeros $0 \cdots 0$, but – as we have shown in Eq. 11 using Eq. 10 – there still exists an asymptotic quasispecies.

How does our simple GA behave with a NiH that moves according to $P_{\oplus\ll}$? In Fig. 6, two typical runs of a genGA with a NiH like that are depicted. The setting (m, l, f_0, τ) was kept fixed but two different mutation rates μ were chosen. In the case of Fig. 6 *(right)*, the mutation rate is 'too high' to allow the population to track the movement. The concentration of the future needle string (solid line) cannot grow much within one jump cycle resulting in a decreasing initial condition (bullet) for the growth of the needle concentration (dotted line) in the next cycle. The population looses the peak – in this case after ≈ 90 generations. It might happen that the population finds the needle again by chance (or better saying the moving needle jumps into the population), but the population will not be able to stably track the movement. Contrasting to that, the mutation rate was chosen to maximize the concentration of the future needle string at the end of each jump cycle (bullets) in Fig. 6 *(left)*. Since in that case, the best achievable initial condition is given to each jump cycle, the movement of the needle



$P_{\oplus\ll}$	
(1) 000→001	(4) 111→110
(2) 001→011	(5) 110→100
(3) 011→111	(6) 100→000
010→101	101→010

Figure 5: The orbit of $0\cdots0$ under $P_{\oplus\ll}$ (black dots) for string length $l = 3$. The numbers $(1), \ldots, (6)$ show the order in which the strings are visited by the needle, starting from 000.

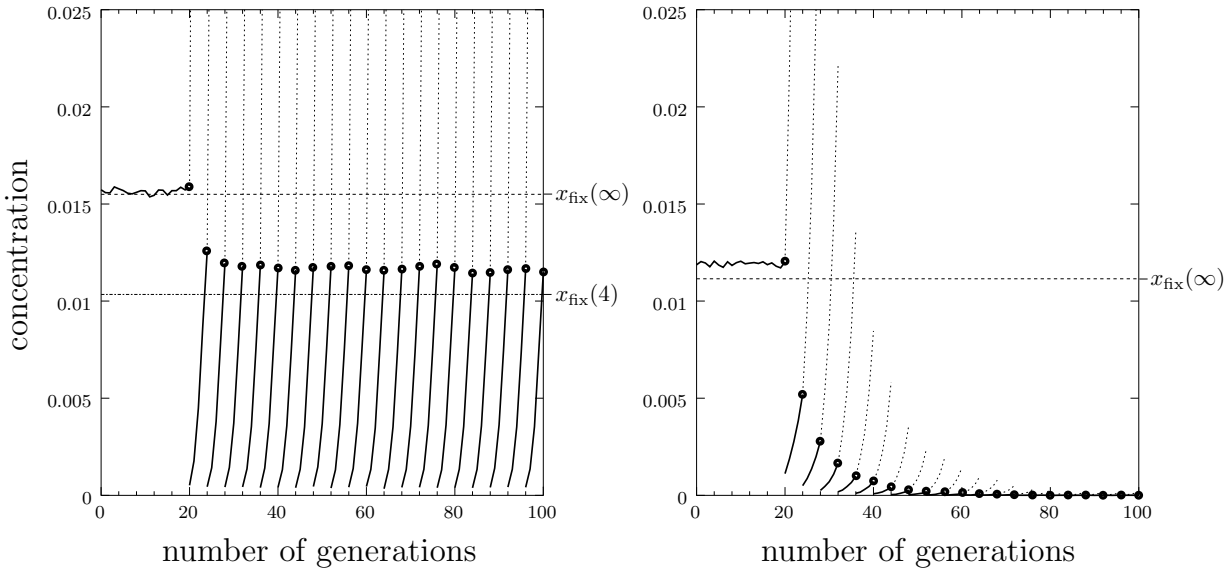


Figure 6: Run of a genGA with *regularly* moving needle-in-the-haystack. The parameter setting was $m = 1000000, l = 20, f_0 = 5, \tau = 4$, *(left):* $\mu = 0.022$, *(right):* $\mu = 0.055$. In both cases the system evolved for 100 generations (not shown) without any occurring jumps in order to let a typical quasispecies grow around the initial needle string. In generation 20 the first jump happened and afterwards every $\tau = 4$ generations. solid line: $x_1(n,t)$, dotted line: $x_0(n,t)$, bullet: jump – $x_0(n+1,0) = x_1(n,\tau)$.

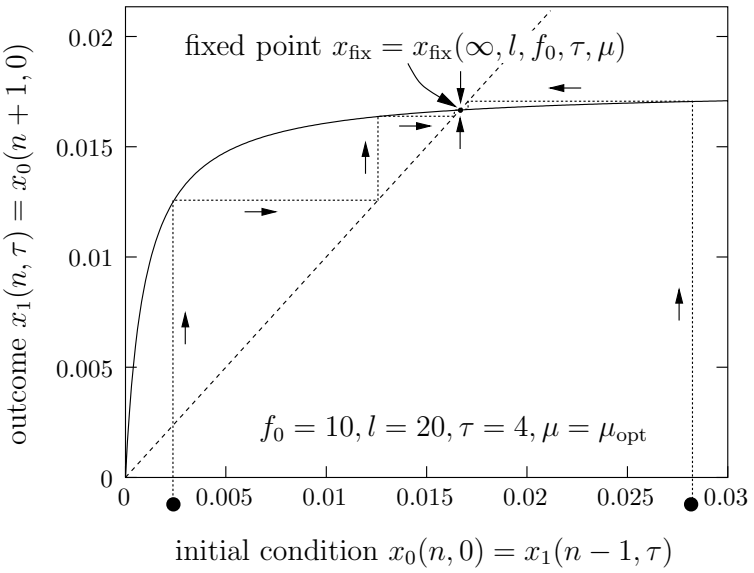


Figure 7: The fixed point which is reached by an infinite population for $n \to \infty$.

is tracked with the highest possible stability for the given setting (m, l, f_0, τ). As can be expected from Fig. 6 and is affirmed by further experiments, the bullets keep on fluctuating around an average value for $n \to \infty$ which is for the infinite population given by the quasispecies Eq. 11. In the following, we are going to model that system with some idealizations and we will calculate a lower boundary for this average value.

We adopt the viewpoint of permuting the concentration vector compatible to the movement of the needle as we have done implicitly in Fig. 6 and formally in the definition of $\mathbf{x}'(t)$ in Eq. 10, but we drop the primes henceforth. The concentration of the needle string within jump cycle n is denoted by $x_0(n, \varphi)$ and the concentration of the string the needle will move to with the $(n+1)$th jump (i. e. the future needle string in jump cycle n) is denoted by $x_1(n, \varphi)$. The initial cycle prior to which *no* jump has occurred is $n = 0$. Within a cycle, the time or generation is counted as phase $\varphi \in \{0, \ldots, \tau\}$. Two succeeding cycles are connected by the (approximated) rule of change

$$x_0(n+1, 0) = x_1(n, \tau) \quad \text{and} \quad x_1(n+1, 0) \approx 0. \tag{14}$$

The second relation is an approximation which is made to simplify the coming calculations, but it holds only if the needle jumps onto a string which has not been close to one of the previous needle positions. Otherwise, the future needle string could already be present with a concentration significantly larger than $1/2^l \approx 0$. In Fig. 6, we have chosen the rule $P_{\oplus\ll}$ to get experimental data for a case in which this assumption is fulfilled. Later on we will see that we can still make useful comments about cases in which that approximation is partly broken.

If we plot $x_0(n+1, 0) = x_1(n, \tau)$ against $x_0(n, 0)$, we get an intuitive picture for the system's evolution towards the quasispecies. The concentration $x_0(n, 0)$ converges

for $n \to \infty$ towards a fixed point,

$$x_{\text{fix}} := \lim_{n\to\infty} x_0(n, 0),$$

as shown in Fig. 7 for a finite value of x_{fix}. Obviously, this fixed point depends on the full setting $x_{\text{fix}} = x_{\text{fix}}(m, l, f_0, \tau, \mu)$. Since we are especially interested in the effects of various cycle lengths τ and mutation rates μ, we keep (m, l, f_0) fixed, such that $x_{\text{fix}} = x_{\text{fix}}(\tau, \mu)$.

In the remaining of this section, we will calculate $x_0(n + 1, 0) = x_1(n, \tau)$ in dependence on $x_0(n, 0)$, which is the solid curve in Fig. 7, for arbitrary parameter settings. From this knowledge, we will construct the phase diagram. Since we stay within one jump cycle, we drop n to take off some notational load.

4.1 Derivation of the Fixed Point Concentrations

To calculate $x_1(\tau)$, it is sufficient to take only x_0 and x_1 into account, because the assumed initial condition is $x_1(0) \approx 0$, such that the main growth of x_1 is produced by the mutational flow from the needle. Moreover, we assume μ to be small enough such that terms proportional to μ^2 can be neglected. This means we restrict ourselves to the case in which the system is mainly driven by one-bit mutations. Without normalization, the evolution equations then read

$$\begin{aligned} y_0(t+1) &= \quad (1-\mu)^l \quad f_0\, y_0(t) + \big\{\mu(1-\mu)^{l-1}\, y_1(t)\big\}, \\ y_1(t+1) &= \mu(1-\mu)^{l-1} f_0\, y_0(t) + \quad (1-\mu)^l \quad y_1(t), \end{aligned} \tag{15}$$

where y_i denote unnormalized concentrations in contrast to the normalized concentrations x_i.

For $f_0(1-\mu) \gg \mu$, which is always the case for large enough f_0, we can further neglect the back-flow $\{\cdots\}$ from the future needle string compared to the self-replication of the current needle string. The solution of Eq. 15 is then given by

$$\begin{aligned} y_0(t) &= \big[(1-\mu)^l f_0\big]^t\, y_0(0), \\ y_1(t) &= \kappa_t(\mu)\, y_0(0) + (1-\mu)^{lt} y_1(0), \end{aligned}$$

$$\text{with } \begin{cases} \kappa_t(\mu) = \mu(1-\mu)^{lt-1}\alpha_t \\ \quad \alpha_t = \sum_{\nu=1}^{t} f_0^{\nu} = f_0 \frac{f_0^t - 1}{f_0 - 1}. \end{cases}$$

The coefficient $\kappa_t(\mu)$ measures the growth of $y_1(t)$ starting from the initial condition $y_1(0) \approx 0, y_0(0) \neq 0$. As long as $y_0(t) + y_1(t) \ll 1$, this gives already a good approximation for the concentrations $x_0(t)$ and $x_1(t)$. But in general, this approximation breaks down for large t, because of the exponential growth of $y_0(t)$. We need to normalize our solution, which can be done by

$$\mathbf{x}(t) = \mathbf{y}(t) \big/ \langle f\rangle_0 \cdots \langle f\rangle_{t-1}, \quad \text{where } \langle f\rangle_t = (f_0 - 1)x_0(t) + 1. \tag{16}$$

By expressing the fitness averages in terms of $y_0(t)$, we find, after solving a simple

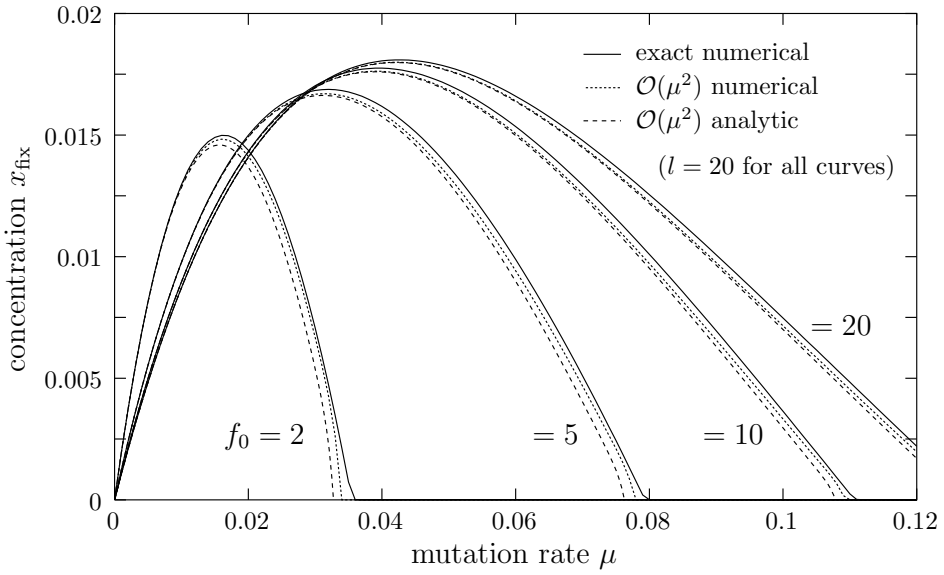


Figure 8: Comparison of the exact numerical and the $\mathcal{O}(\mu^2)$ calculation for different values of the needle fitness f_0.

recursion,

$$\begin{aligned}\langle f\rangle_0 \cdots \langle f\rangle_{t-1} &= 1 + (f_0 - 1)\left[\sum_{\nu=0}^{t-1}(1-\mu)^{l\nu} f_0^{\nu}\right] x_0(0) \\ &= 1 + (f_0 - 1)\beta_t(\mu) x_0(0),\end{aligned}$$

$$\text{where } \beta_t(\mu) = \tfrac{\tilde{f}^t - 1}{\tilde{f} - 1} \text{ and } \tilde{f} = (1-\mu)^l f_0.$$

Finally, we arrive at the normalized concentrations

$$\begin{aligned}x_0(t) &= \left[(1-\mu)^l f_0\right]^t x_0(0) \Big/ \left[1 + (f_0 - 1)\beta_t(\mu) x_0(0)\right], \\ x_1(t) &= \left[\kappa_t(\mu)\, x_0(0) + (1-\mu)^{lt} x_1(0)\right] \Big/ \left[1 + (f_0 - 1)\beta_t(\mu) x_0(0)\right].\end{aligned}$$

The asymptotic state can now be calculated by using the initial condition $x_1(0) \approx 0, x_0(0) \neq 0$ and demanding $x_1(\tau) = x_0(0)$. It is easily verified that for the fixed point follows

$$x_{\rm fix}(\tau, \mu) = \frac{\kappa_\tau(\mu) - 1}{(f_0 - 1)\beta_\tau(\mu)}. \tag{17}$$

4.2 Consistency in the Quasi-Static Limit

How can we test the quality of the approximate result Eq. 17? For large cycle lengths τ, we enter the quasi-static regime, where we can approximate the population at the end of each cycle by the quasispecies of the corresponding static landscape. Figure 8 shows a comparison of the exact numerical calculations of the quasispecies ($\tau \to \infty$) and the $\mathcal{O}(\mu^2)$ calculations ($\tau = 100$). In the numerical $\mathcal{O}(\mu^2)$ calculation, the back-flow from the first error class to the needle string is included. Overall, we find the error

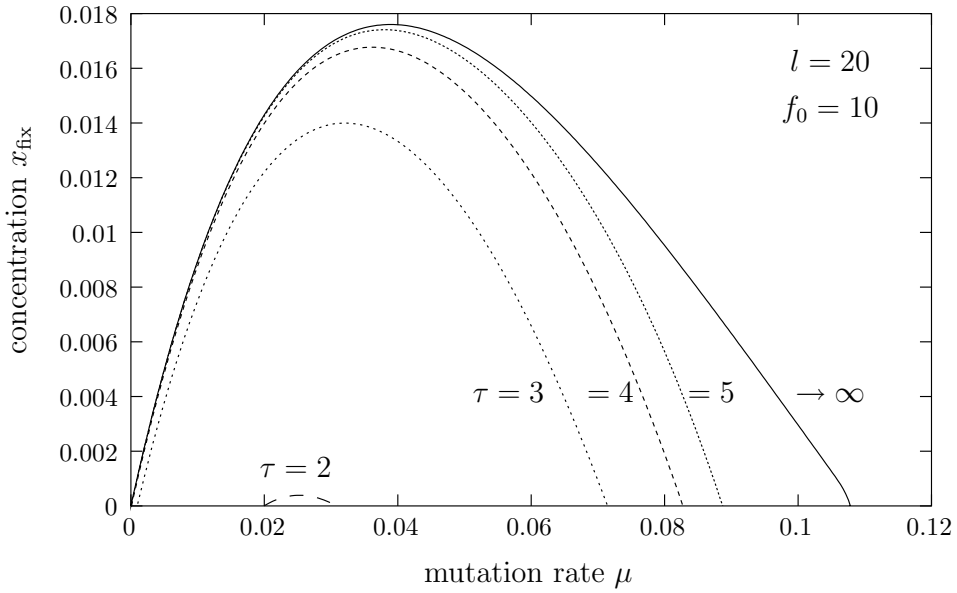


Figure 9: Fixed point concentration $x_{\text{fix}}(\tau, \mu)$ for different values of τ. For faster changes, the fixed point concentration rapidly drops down.

threshold and the maximum of the fixed point concentration well represented. This also suggests that the deviation of the $\mathcal{O}(\mu^2)$ approximation from the exact values should be small for smaller τ, because those deviations add up for $\tau \to \infty$ by the iterative procedure.

How do the calculated fixed point concentrations compare to simulations with (large) finite population? In Fig. 6, the values of $x_{\text{fix}}(\infty, \mu)$ and $x_{\text{fix}}(4, \mu)$ are shown. For $\tau \to \infty$, the deviation from the average $\langle x_1(n, \varphi) \rangle$ (in generations $0 - 20$) is in fact the same as what can be read off in Fig. 8. The deviation of $x_{\text{fix}}(4, \mu)$ from the average value $\langle x_0(n, 0) \rangle$ in generations $24, 28, \ldots, 100$ is significantly larger. This is caused by the neglect of all other strings' contributions apart from the current needle string's contribution to the flow onto the future needle string. These neglected contributions increase the average fixed point concentration measured in the experiment in comparison to the calculated value $x_{\text{fix}}(\tau, \mu)$. But even though there are deviations, we conclude that the approximately calculated value is always a lower bound for the exact value. In the next section, we will use this observation to derive an expression for the mutation rate that maximizes the average fixed point concentration.

4.3 Phase Diagram

In Fig. 9, the fixed point values $x_{\text{fix}}(\tau, \mu)$ are shown for small cycle lengths τ. For the shown parameter setting, the region with $x_{\text{fix}}(2, \mu) > 0$ is extremely small. We notice that there are two error thresholds, one for 'too low' mutation rates, $\mu_{\text{th}<}$, and one for 'too high' mutation rates, $\mu_{\text{th}>}$. The intuition behind that was already given in Section 3. For too low mutation rates the population becomes slow and evolves in the averaged, flat landscape, whereas for too high mutation rates the usual transition to the disordered phase takes place. In the following we will calculate the phase diagram starting from Eq. 17.

Error Thresholds: The error thresholds are given by

$$x_{\text{fix}}(\tau, \mu) = 0 \quad \Longleftrightarrow \quad \kappa_\tau(\mu) = 1. \tag{18}$$

This is the same condition as one would get using only unnormalized concentrations $y_i(t)$. Since $y_i(t) \approx 0$ near the error thresholds, the neglect of the normalization is not critical for the calculation of the error thresholds themselves, whereas it is important for the optimal mutation rate and of course for the fixed point concentration. Since Eq. 18 cannot be solved for μ in closed form, we write down the corresponding recursion relation that converges, for a suitable starting value of μ, to the solution of Eq. 18 in the limit $k \to \infty$,

$$\begin{aligned}
\mu_{\text{th}<}^{(k)} &= 1 \Big/ \alpha_\tau \left(1 - \mu_{\text{th}<}^{(k-1)}\right), & \mu_{\text{th}<}^{(0)} &= 0, \\
\mu_{\text{th}>}^{(k)} &= 1 - \left(1 \Big/ \alpha_\tau \mu_{\text{th}>}^{(k-1)}\right)^{1/(l\tau-1)}, & \mu_{\text{th}>}^{(0)} &= 1 - f_0^{-1/l} =: \mu_{\text{th}}^{\infty}.
\end{aligned}$$

For $\mu_{\text{th}<}$, a good starting value is 0, since $\mu_{\text{th}<} \approx 0$ anyway. For $\mu_{\text{th}>}$, the approximate value for the error threshold of the static (i. e. $\tau \to \infty$) landscape μ_{th}^{∞} can be chosen, which is obtained by calculating the fixed point [using Eq. 15 and 16],

$$x_0(t+1) = x_0(t) \quad \Longleftrightarrow \quad x_{\text{fix}}^{\infty} = \frac{(1-\mu)^l f_0 - 1}{f_0 - 1},$$

setting it to zero and solving for μ.

Optimal Mutation Rate: In order to track changes with the best achievable stability for a given setting (m, l, f_0, τ), the lowest possible concentration (infimum of) $x_0(n, \varphi)$ needs to be maximized, because a low concentration might result in the loss of the needle string in a finite population. Since for infinite populations $x_0(n, \varphi)$ is monotonously increasing with φ it is sufficient to maximize $x_0(n, 0)$. Moreover, we derived above that $x_0(n, 0)$ approaches the fixed point value $x_{\text{fix}}(\tau, \mu)$ for $n \to \infty$. For finite populations, we expect similar behavior but the strict monotony of $x_0(x, \varphi)$ in φ will be destroyed by fluctuations and also the fixed point value itself will fluctuate around some average value $\langle x_{\text{fix}} \rangle$ as can be seen in Fig. 6. However, the *safest* way to avoid any loss of the needle string is still to maximize the average fixed point value $\langle x_{\text{fix}} \rangle$. In this sense, we define the *optimal mutation rate* μ_{opt} as the one that maximizes $\langle x_{\text{fix}} \rangle$. In the previous Section 4.2, we noted that our approximated infinite population value $x_{\text{fix}}(\tau, \mu)$ represents a lower bound for $\langle x_{\text{fix}} \rangle$, where the maxima of the two curves are expected to coincide for fixed τ. Thus, μ_{opt} can be obtained by maximization of $x_{\text{fix}}(\tau, \mu)$.

We can derive an expression for the optimal mutation rate μ_{opt} from

$$\frac{\partial x_{\text{fix}}}{\partial \mu}(\tau, \mu_{\text{opt}}) = 0$$

If we neglect the μ dependence of $\beta_\tau(\mu)$ in Eq. 17, which corresponds to the approach in [19], we simply find $\mu_{\text{opt}}^{\text{NS}}(\tau, l) = 1/l\tau$. Because of $\mu_{\text{opt}}^{\text{NS}} \xrightarrow{\tau \to \infty} 0$, this result is inconsistent with the quasi-static limit, because μ_{opt} should approach the value for which the

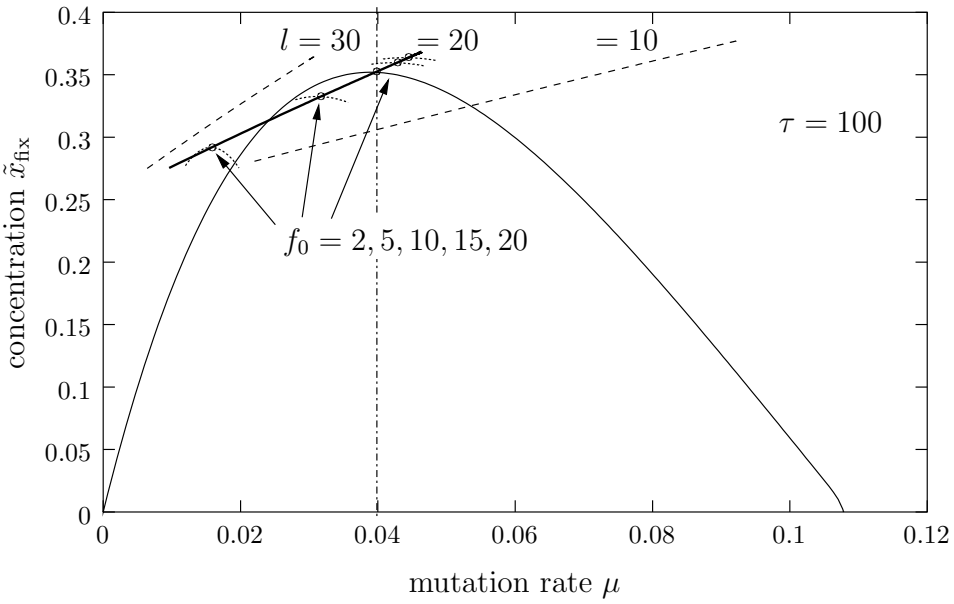


Figure 10: The optimal mutation rate $\mu_{\text{opt}}^{\infty}(f_0, l)$ from Eq. 19 in dependence on needle height f_0 and string length l.

concentration of 1-mutants in the quasispecies of the corresponding static NiH landscape is maximized. We conclude that the μ dependence of $\beta_\tau(\mu)$ cannot be neglected for the correct optimal mutation rate, which we are now going to calculate.

For $\alpha_\tau \gg 1$, which is the case for $\tau \gg 1$ and $f_0 > 1$, or $\tau \approx 1$ and $f_0 \gg 1$, we can neglect the -1 in the numerator of $x_{\text{fix}}(\tau, \mu)$ and take only α_τ into account for the calculation of $\partial x_{\text{fix}}/\partial\mu$. After some algebra, we find

$$\mu_{\text{opt}} = \frac{(\tilde{f}^\tau - 1)(\tilde{f} - 1)}{l(\tilde{f}^{\tau+1} - (\tau+1)\tilde{f} + \tau)}, \quad \text{where } \tilde{f} = f_0(1-\mu_{\text{opt}})^l.$$

Since $\tilde{f} = \tilde{f}(\mu_{\text{opt}})$, this equation cannot be solved in a closed form for μ_{opt}. However, for $\tau \to \infty$ the equation simplifies to

$$\mu_{\text{opt}}^{\infty} = \begin{cases} (\tilde{f} - 1)/l\tilde{f} & : \tilde{f} > 1 \\ 0 & : \tilde{f} \leq 1. \end{cases}$$

In the case $\tilde{f} > 1$, we find

$$(1 - l\mu_{\text{opt}}^{\infty})(1 - \mu_{\text{opt}}^{\infty})^l = 1/f_0.$$

By approximating $(1-\mu)^l \approx (1-l\mu)^2$, we get a cubic equation. The real root of that equation is approximately [20] given by (see also Fig. 10)

$$\mu_{\text{opt}}^{\infty}(f_0, l) \approx \mu_+ \left[1 + \frac{(l-1)\mu_+(1-l\mu_+)}{3l(l-1)\mu_+^2 - 2\mu_+(3l-1) + 4}\right]$$

$$\text{with } \mu_+ = \frac{1}{l}\left[1 + f_0^{-1/2}\right]. \quad (19)$$

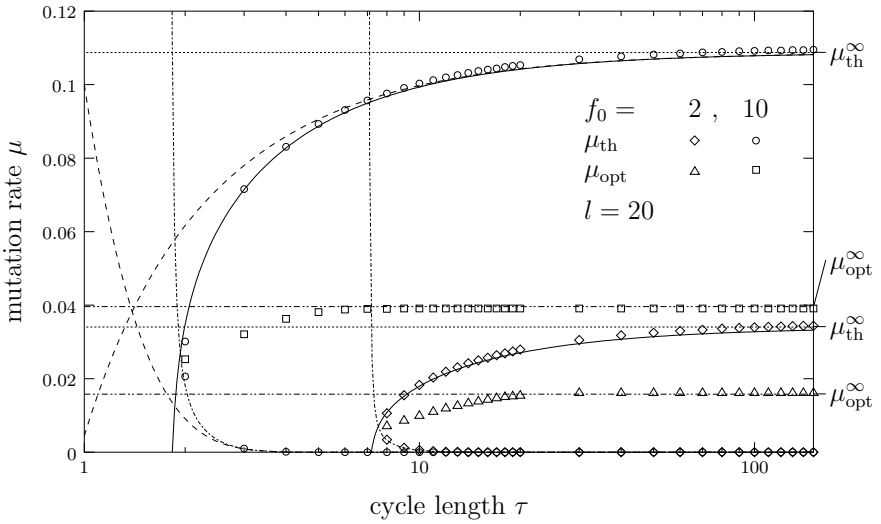


Figure 11: The calculated phase diagram for a genGA with stochastically moving needle-in-the-haystack; two settings are shown: $f_0 = 2, 10$, for both $l = 20$.

Resulting Phase Diagram: From the above, we are able to plot the phase diagram for our model as shown in Fig. 11. Two settings are plotted. For $f_0 = 2$ (resp. 10) the diamonds (resp. circles) are the numerically obtained error thresholds. The solid and dash-dotted lines are $\mu^{(5)}_{\rm th<}$ and $\mu^{(5)}_{\rm th>}$. To show the convergence property of $\mu^{(k)}_{\rm th<,>}$, $\mu^{(0)}_{\rm th<,>}$ are plotted for $f_0 = 10$ as dashed lines. Obviously, the needed corrections to the chosen starting values increase for smaller τ, such that more iterations are needed to describe the error thresholds correctly for small τ. The expressions $\mu^{(5)}_{\rm th<,>}$ are already a good approximation for the given settings. Representing the quasi-static limit, $\mu^{\infty}_{\rm th}$ is plotted as dotted line and gets consistently approached by $\mu_{\rm th>}(\tau)$ for $\tau \to \infty$. Furthermore, $\mu^{\infty}_{\rm opt}$ is plotted as dash-dot-dotted line. The numerically measured values for $\mu_{\rm opt}(\tau)$ are shown for $f_0 = 2$ (resp. 10) as triangle (resp. squares). They approach $\mu^{\infty}_{\rm opt}$ very quickly already for $\tau \approx 20$ (resp. 10).

We conclude that the above quantitative description is in good agreement with the numerical observations and approaches the quasi-static region in a consistent way. Moreover, the phase diagram fits well into the general one raised in Section 3. Even in the considered case of a genGA, we find – depending on the parameter setting – a time-averaged phase for very small τ. The time-averaged phase broadens for small f_0.

4.4 Stochastically moving NiH

Up to now, we analyzed a regularly moving NiH, for example with the rule $P_{\oplus\ll}$. What happens if the NiH is allowed to move to a *randomly* picked nearest neighbor, as it is shown in Fig. 12 for $l = 4$? Two typical runs of a genGA with this fitness landscape are depicted in Fig. 13. The setting (m, l, f_0, τ) was chosen the same as in Fig. 6 which allows for a direct comparison of the GA's behavior for regularly and stochastically moving NiHs. The overall behavior is similar. For large mutation rates, the population looses the needle string, whereas the moving needle is tracked stably for mutation rates close to the above defined optimal mutation rate. In addition, strong fluctuations in

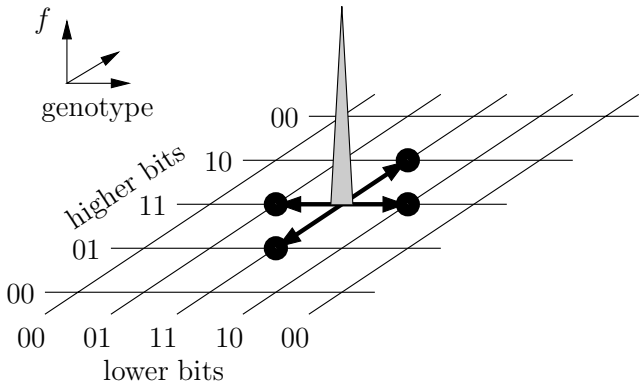


Figure 12: A *stochastically* moving needle-in-the-haystack for string length $l = 4$. The needle is allowed to jump to one of its nearest neighbors which is chosen at random.

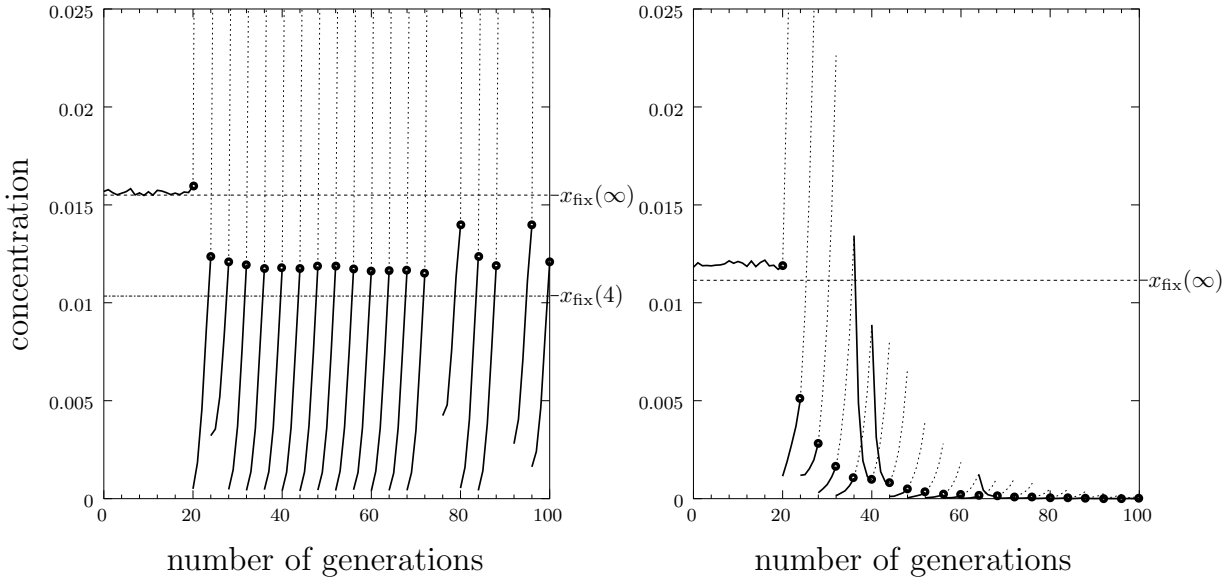


Figure 13: Run of a genGA with *stochastically* moving needle-in-the-haystack. The parameter setting in *(left)* and *(right)* were the same as in Fig. 6 *(left)* and *(right)*.

the values of $x_1(n,0)$ (lower ends of solid lines) as well as $x_0(n+1,0) = x_1(n,\tau)$ (bullets) occur in the *stochastic* case. These result from *back-jumps*. If, at the end of the current cycle, the needle jumps back to the string it has been to in the previous cycle, then $x_1(n,0) = x_0(n-1,\tau)$ is significantly larger than zero. This can be seen in Fig. 13 *(right)* at generations $36, 40$ and 64 and also in Fig. 13 *(left)* at generations 72 and 88 (the gaps in Fig. 13 *(left)* correspond to x_1, x_0 being much larger than 0.025). If no back-jumps occur, as in generations $24 - 72$ in Fig. 13 *(left)*, the system with stochastic NiH behaves nearly indistinguishable from the one with regularly moving NiH. Since back-jumps always increase the concentrations of the needle string in the very next occurring jumps, the above calculated fixed point $x_{\text{fix}}(\tau,\mu)$ is still a lower bound. Thus, our previous notion of optimal mutation rate remains applicable to the stochastically moving NiH although the assumption $x_1(n,0) \approx 0$ from Eq. 14 is not always fulfilled.

Nilsson and Snoad [19] did their analysis of the continuous Eigen model Eq. 3 with stochastic NiH in a similar way as we did above. In analogy to their calculation for the continuous Eigen model, we find for a genGA the optimal mutation rate $\mu_{\text{opt}}^{\text{NS}}(\tau,l) = 1/l\tau$ which is inconsistent with the quasi-static limit (see Section 4.3). The reason is the missing normalization in the work of Nilsson and Snoad. Furthermore, they could not derive an expression for the fixed point concentration $x_{\text{fix}}(\tau,\mu)$ because of that same reason.

4.5 Jumps of larger Distance

To conclude this section about the behavior of genGAs with different kinds of NiHs that move to *nearest* neighbors, let us shortly discuss jumps of Hamming distance d *larger* than one. Obviously, the analytical calculations get more complicated, because the $\mathcal{O}(\mu^2)$-approximation is not sufficient anymore as it connects only nearest neighbors. To describe jumps of a larger distance, the concentrations of some intermediate sequences need to be taken into account, so that we have to solve a time evolution much more complicated than Eq. 15. Hence, we cannot make simple statements for finite τ. On the other hand, the system approaches the quasi-static region for large τ and it is characterized by $\mu_{\text{th}<,>}^{\infty}$ and $\mu_{\text{opt}}^{\infty}$ as we have seen in Fig. 11. The exact quasispecies for $\tau \to \infty$ is shown in Fig. 14. The plotted values are error class concentrations, in order to make the higher error classes visible at all. Each k-mutant has a concentration of $\tilde{x}_k/\binom{l}{k}$ in the quasispecies state, because for a NiH the mutant's fitness depends only on its Hamming distance to the needle and therefore all $\binom{l}{k}$ k-mutants have the same concentration in the quasispecies. For finite populations, this is only true on average, because the asymptotic state is distorted by fluctuations. But in the following, we assume that the quasispecies is still representative for the average distribution of the population in the asymptotic state. Then, the optimal mutation rate in the sense of Section 4.3 for jumps of distance d is by definition the position of the maximum of $\tilde{x}_d$. For $d \geq l/2$, optimal mutation rate and error threshold become identical. Although $\tilde{x}_d$ is maximized for mutation rates close to the error threshold it amounts, as do all other concentrations to only $\approx 1/2^l$, which leads to an approximately random drift for finite populations. On the other hand, the chance of tracking the needle decreases even further for small mutation rates because then the concentration $\tilde{x}_d$ becomes even smaller. In this sense, the quasispecies distribution, which is centered on the needle string, is

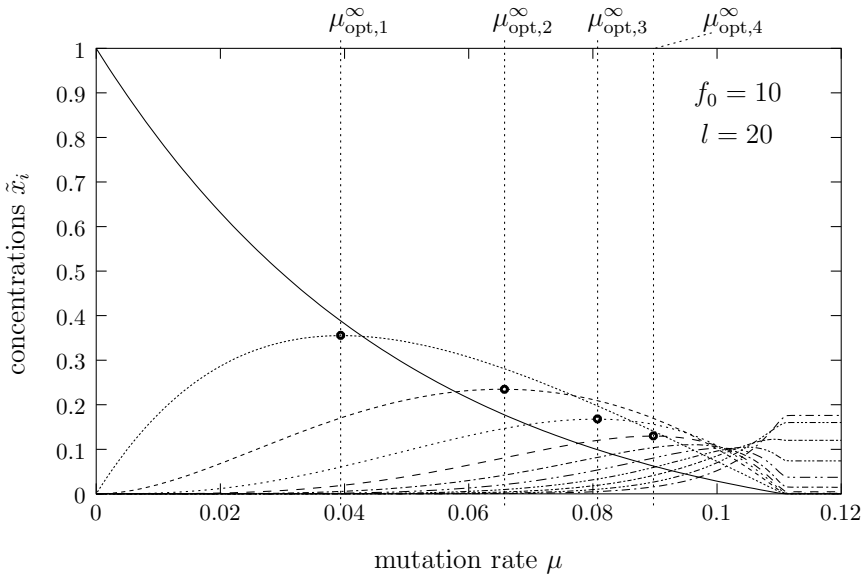


Figure 14: The quasispecies for the static NiH in dependence on the mutation rate μ. The concentrations $\tilde{x}_i$ of the ith error class for $i \in \{0, \ldots, \lfloor l/2 \rfloor\}$ are depicted. The optimal mutation rates for jumps of Hamming distance $d = 1, 2, 3, 4$ are shown as dotted lines.

useless for tracking the next jump if $d \geq l/2$. This also suggests – in agreement with the experimental findings of Rowe [13] (in this book) – that finite populations are for low mutation rates unable to track large jumps – in particular in the extreme case $d = l$. Only for jumps of $d < l/2$ the corresponding error class concentration $\tilde{x}_d$ shows a concentration maximum significantly above $1/2^l$. From the heights of the concentration maxima, we see that the difficulty of tracking the changes increases with the Hamming distance d of the jumps. Vice versa, the advantage a population gets after a jump from its structure prior to the jump decreases with increasing jump distance d. In addition, a mutation rate which is simultaneously optimal for more than one distance cannot be found.

5 Conclusions and Future Work

On the basis of general arguments, the phase diagrams of population-based mutation and probabilistic selection systems like the above genGA, ssGA and Eigen model in time-dependent fitness landscape can be easily understood. The notion of regular changes allows for an exact calculation of the asymptotic state in the sense of a generalized, time-dependent quasispecies. For a genGA with NiH that moves regularly to nearest neighbors, the quasispecies can be straightforwardly calculated under simplifying assumptions. The result is a lower bound for the exact quasispecies. With that lower bound, we have constructed the phase diagram in the infinite population limit. This phase diagram is in agreement with the one raised from general arguments.

In order to improve our analysis, we need to weaken our assumptions. In particular, we have to overcome the restriction of taking into account only the flow from the current towards the future needle string. The presence of other contributions to the

flow has to be modeled in some way. Another future step could be an investigation of the fluctuations that are introduced by the finiteness of realistic populations (discreteness of Λ_m) around the quasispecies. This would lead to a lower boundary for the population size above which the needle string is not lost due to those fluctuations.

An extension of our analysis to non-regularities like the occurrence of more than a single jump rule, can be achieved by averaging the time evolution Eq. 8 for $n \to \infty$ according to each rule's probability of being applied. A similar averaging procedure will be necessary if fluctuations of the cycle length τ are present. Finally, an extension of the description to broader, more realistic peaks, as well as GA models including crossover and other selection schemes, are important topics for future work.

References

[1] T. Bäck, U. Hammel and H.-P. Schwefel. *Evolutionary Computation: Comments on the History and Current State*. IEEE Transactions on Evol. Comp. 1(1), p. 3, 1997.

[2] T. Bäck, D. B. Fogel and Z. Michalewicz, editors. *Handbook of Evolutionary Computation*. IOP Publishing, Bristol, 1997.

[3] J. Branke. *Evolutionary Algorithms for Dynamic Optimization Problems, A Survey*. Technical Report 387, AIFB University Karlsruhe, 1999.

[4] J. E. Rowe. *Finding attractors for periodic fitness functions*. In W. Banzhaf *et al.*, editors, *Proceedings to GECCO 1999*, Morgan Kaufmann, San Mateo, p. 557, 1999.

[5] L. M. Schmitt, C. L. Nehaniv and R. H. Fujii. *Linear analysis of genetic algorithms*. Theoretical Computer Science 200, p. 101, 1998.

[6] M. Eigen. *Selforganization of matter and the evolution of biological macromolecules*. Naturwissenschaften 58, p. 465, 1971.

[7] M. Eigen and P. Schuster. *The Hypercycle – A Principle of Natural Self-Organization*. Springer-Verlag, Berlin, 1979.

[8] M. Eigen, J. McCaskill and P. Schuster. *The molecular quasispecies*. Adv. Chem. Phys. 75, p. 149, 1989.

[9] E. Baake and W. Gabriel. *Biological evolution through mutation, selection, and drift: An introductory review*. Ann. Rev. Comp. Phys. 7, in press, 1999.

[10] J. E. Rowe. *The dynamical systems model of the simple Genetic Algorithm*. this issue, p. XXX, 1999.

[11] M. D. Vose. *The simple Genetic Algorithm – Foundations and Theory*. MIT Press, Cambridge, 1999.

[12] E. van Nimwegen, J. P. Crutchfield and M. Mitchell. *Statistical Dynamics of the Royal-Road genetic algorithms*. Theoretical Computer Science, special issue on Evolutionary Computation, A. Eiben, G. Rudolph, editors, in press, 1998.

[13] J. E. Rowe. *Cyclic Attractors and Quasispecies Adaptability*. this issue, p. XXX, 1999.

[14] K. DeJong and J. Sarma. *Generation Gaps Revisited*. In L. D. Whitley, editor, *Foundations of Genetic Algorithms 2*, Morgan Kaufmann, San Mateo, p. 19, 1993.

[15] A. Rogers and A. Prügel-Bennett. *Modeling the Dynamics of a Steady State Genetic Algorithm*. In W. Banzhaf and C. Reeves, editors, *Foundations of Genetic Algorithms 5*, Morgan Kaufmann, San Mateo, p. 57, 1998.

[16] J. Branke, M. Cutaia and H. Dold. *Reducing Genetic Drift in Steady State Evolutionary Algorithms*. In W. Banzhaf *et al.*, editors, *Proceedings to GECCO 1999*, Morgan Kaufmann, San Mateo, p. 68, 1999.

[17] C. O. Wilke, C. Ronnewinkel and T. Martinetz. *Molecular Evolution in time-dependent Environments*. In D. Floreano, J.-D. Nicoud and F. Mondada, editors, *Proceedings to European Conference on Artificial Life 1999*, Springer, Berlin, p. 417, 1999.

[18] C. O. Wilke and C. Ronnewinkel. *Dynamic Fitness landscapes in the Quasispecies model*. in preparation.

[19] M. Nilsson and N. Snoad. *Error Thresholds on dynamic Fitness-Landscapes*. Working Paper 99-04-030, Santa Fe Institute, 1999.

[20] A more detailed explanation and analysis of the used approximation will be presented elsewhere.