PHYSICAL REVIEW E, VOLUME 65, 031901

Quasispecies evolution on a fitness landscape with a fluctuating peak

Martin Nilsson*
Institute of Theoretical Physics, Chalmers University of Technology and Göteborg University, S-412 96 Göteborg, Sweden

Nigel Snoad†
Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501
and The Australian National University, ACT 0200 Canberra, Australia

(Received 19 April 2000; revised manuscript received 28 September 2001; published 7 February 2002)

A quasispecies evolving on a fitness landscape with a single peak of fluctuating height is studied. In the approximation that back mutations can be ignored, the rate equations can be solved analytically. It is shown that the error threshold on this class of dynamic landscapes is defined by the time average of the selection pressure. In the case of a periodically fluctuating fitness peak, we also study the phase shift and response amplitude of the previously documented low-pass filter effect. The special case of a small harmonic fluctuation is treated analytically.

DOI: 10.1103/PhysRevE.65.031901 PACS number(s): 87.23.Kg, 87.10.+e, 87.16.Ac

I. INTRODUCTION

A system consisting of self-replicating genomes can be studied using the quasispecies model [1]. The evolutionary dynamics is driven by point mutations and selection defined by an underlying fitness landscape. A quasispecies is then an equilibrium distribution of closely related gene sequences, localized around one or a few sequences of high fitness. In the case of a static fitness landscape the kinetics of these simple systems has been studied in great detail; see for instance Refs. [1–8].

The main result derived from the model is the appearance of an error threshold (or error catastrophe) as an upper bound on the mutation rate above which no effective selection can occur. This places limits on the maintainable amounts of genetic information [1,2] in simple biological systems.

Until now studies of quasispecies have mainly focused on static fitness landscapes. The main reason for this is perhaps the rather severe mathematical difficulties appearing when the fitness landscape is allowed to be time dependent (these difficulties will be discussed in more detail later). However, many organisms in nature live in a quickly changing environment. In this paper, we show that a simple approximation allow the population dynamics to be analyzed analytically when the fitness landscape consists of a fluctuating single peak. The expression for the error threshold is obtained from the expression in the static case by replacing the height of the fitness peak by the time average of the height of the fluctuating peak. We also study how the phase shift between fitness oscillations and population dynamics depends on the frequency in the case of a small harmonic fluctuation.

II. QUASISPECIES IN DYNAMIC ENVIRONMENTS

A quasispecies consists of a population of self-replicating genomes, where each genome is represented by a sequence of bases s_k, $(s_1 s_2 \cdots s_\nu)$. We assume that the bases are binary, $s_k \in \{1,0\}$ and that all sequences have equal length ν. Every genome is then given by a binary string $(011001\cdots)$, which also can be represented by an integer $k = \Sigma_j s_j 2^j$ $(0 \leqslant k < 2^\nu)$.

To describe how mutations affect a population, we define W_k^l as the probability that replication of genome l gives genome k as offspring. We only consider point mutations, which conserve the genome length ν.

We assume that the point mutation rate $\mu = 1 - q$ (where q is the copying accuracy per base) is constant in time and independent of the position in the genome. We can then write down an explicit expression for W_k^l in terms of the copying fidelity:

$$W_k^l = \mu^{h_{kl}} q^{\nu - h_{kl}} = q^\nu \left(\frac{1-q}{q} \right)^{h_{kl}}, \tag{1}$$

where h_{kl} is the Hamming distance between genomes k and l. The Hamming distance h_{kl} is defined as the number of positions where the genomes k and l differ.

The equations describing the dynamics of the population take a relatively simple form. Let x_k denote the relative concentration and $A_k(t)$ the time-dependent fitness of genome k. We then obtain the following rate equations:

$$\dot{x}_k(t) = \sum_l W_k^l A_l(t) x_l(t) - e(t) x_k(t), \tag{2}$$

where $e(t) = \Sigma_l A_l(t) x_l(t)$, and the dot denotes a time derivative. The second term ensures the total normalization of the population $[\Sigma_l x_l(t) = 1]$ so that $x_k(t)$ describe relative concentrations.

Equation (2) can be linearized by a change of variable, $y_k(t) = e^{\int_{t_0}^{t} e(s) ds} x_k(t)$. Giving a set of linear differential equations

$$\dot{y}_k(t) = \sum_l W_k^l A_l(t) y_l(t). \tag{3}$$

---

*Email address: martin@fy.chalmers.se

†Email address: nigel@santafe.edu

In the classical theory introduced by Eigen and coworkers [1,9,10], the fitness landscape is static. The rate equations (2) can then be solved analytically by introducing a change of coordinates that makes them linear, and then solving the eigenvalue system for the matrix $W_k^l A_l$. The equilibrium distribution is given by the eigenvector corresponding to the largest eigenvalue.

When the fitness landscape is time dependent, this method cannot be applied. A time-ordering problem occurs when we define exponentials of time-dependent matrices, since in general the matrix $W_k^l A_l(t)$ does not commute with itself at different points in time. Formally we may describe the dynamics of the system by a *time-ordered exponent*

$$\begin{aligned} y(t) &= \exp_+\left\{\int_{t0}^{t}\hat{O}(s)ds\right\}y_0 \\ &= \sum_{n=0}^{\infty}\int_{t_0}^{t}dt_n\cdots\int_{t_0}^{t_2}dt_1\,\hat{O}(t_n)\cdots\hat{O}(t_1)y_0 \\ &= \sum_{n=0}^{\infty}\frac{1}{n!}\int_{t_0}^{t}dt_n\cdots\int_{t_0}^{t}dt_1\,T[\hat{O}(t_1)\cdots\hat{O}(t_n)]y_0, \end{aligned} \tag{4}$$

where $\hat{O}(t)$ is the operator notation for the matrix $W_k^l A_l(t)$ and the symbol $T[\cdot]$ shows that the operators $\hat{O}(t_i)$ are ordered from the right to the left as time increases. Equation (5) is, however, mainly of formal use and does not directly increase our understanding of the dynamics.

For a more direct approach, we may consider using the relation

$$e^X e^Y = e^{H(X,Y)}, \tag{5}$$

where $H(X,Y)$ is a Lie series (sometimes referred to as the *Hausdorff series*), i.e., a series whose terms are elements in the Lie algebra generated by X and Y. For general (noncommuting) the first few terms read

$$H(X,Y) = X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}([X,[X,Y]] + [[X,Y],Y]) + \cdots. \tag{6}$$

To the best of our knowledge, there is no reasonable explicit expression for all the terms in the Hausdorff series. Just like Eq. (5), it is, therefore, only of limited use when analyzing the linearized rate equations on dynamic fitness landscapes.

Later in this paper, we make a simple approximation that makes the rate equations one dimensional; time ordering is then no longer necessary and analytical solutions are relatively easy to derive.

Much of the work on quasispecies has focused on fitness landscapes with one gene sequence (the master sequence) with superior fitness σ compared to all other sequences. These are viewed as a background with fitness 1. These landscapes are referred to as single peaked landscapes. The main reasons why this particular fitness landscape has attracted much attention is because it is the simplest landscape exhibiting an error threshold and it is also possible to calculate a simple algebraic relation for the threshold [2]. On this landscapes, the master sequence is denoted x_0.

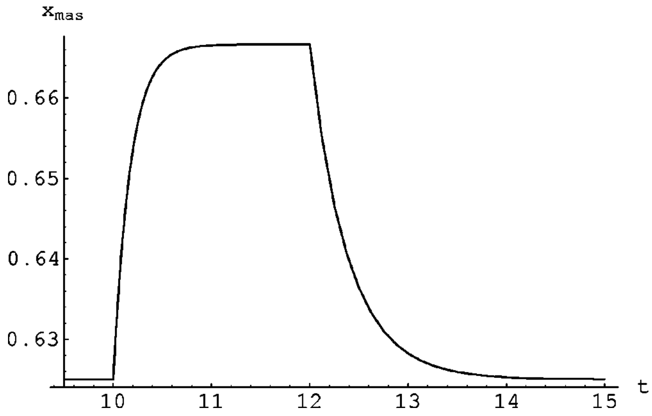

FIG. 1. Concentration of the master sequence when the fitness peak makes a sudden jump. The fitness is given by $\sigma = 10$ when $t \in [10,12]$, $\sigma = 5$ otherwise. The genomic copying fidelity is given by $Q = 0.7$.

In a previous paper [17], we studied quasispecies evolution on a fitness landscape with a single "wandering" peak, i.e., the fitness peak where static during predefined time periods and then moved instantly to a new location in the neighborhood of the old peak. In this case, we may solve the rate equations during the static periods by the standard methods and then match these solutions together to find a global solution. In this paper, we focus on single peaked landscapes where the height of the fitness peak is time dependent. The fitness landscape is then given by

$$A_k(t) = \begin{cases} \sigma(t) & \text{if } k = 0 \\ 1 & \text{otherwise} \end{cases}. \tag{7}$$

This class of time-dependent landscapes has been studied extensively by Wilke and co-workers [11,12]. They investigated the behavior of a periodically fluctuating single peak landscape by numerically integrating the dynamics to find the limit cycle of the concentrations for a full period. They also find that the population response can be described as a low-pass filter (this effect had already been pointed out by others, see e.g., Refs. [14–16,18]). Wilke, Ronnewinkel, and Martinetz also studied how the phase shift between the fitness landscape and the population response depend on the fluctuation frequency. Further they numerically demonstrate that for rapidly fluctuating landscapes, the error threshold occurs when the time average of the fitness falls below the corresponding static threshold. In this paper, we demonstrate that all these effects can be studied analytically using a simple approximation.

Figure 1 shows how the concentration of the master sequence responds to a sudden, sharp jump in its fitness. When the fitness changes it takes some time for the population to reach the new equilibrium. It is this delay that causes a phase shift between a periodically changing fitness function and the response in the concentrations. The relaxation time of the population to the appropriate equilibrium distribution de-

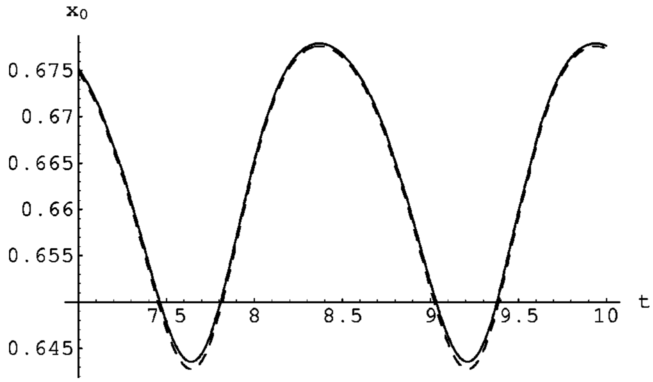

FIG. 2. Concentration of the master sequence calculated using the approximation 8 (dashed) and by numerically solving the rate equations (2) (solid). The fitness is given by $\sigma = 10 + 5\ \sin(4t)$. The genomic copying fidelity is given by $Q = 0.7$ and the genome length $\nu = 25$.

pends on both the fitness values of the landscape and the mutation rate. For extremely slow and smooth changes in the fitness the population will effectively reach equilibrium at every point in time. Thus the continued existence of a quasispecies will depend on the local dynamics of the landscape. When the landscape changes quickly the population will fail to follow the changes adequately and thus responds to the landscape dynamics in a way that is typical of a low-pass filter. The following section examines the fluctuating single peak landscape in some detail. In particular, we introduce an approximation that lets us find an analytic form for the relaxation time of the population, and the phase lag it introduces in a periodic landscape.

III. APPROXIMATE QUASISPECIES DYNAMICS

We now introduce a simple approximation of the model presented above. In this approximation, we can solve the rate equations and find an expression for the concentration of the master sequence $x_0(t)$. In the limit of long chain length ($\nu \gg 1$), we can neglect back mutations from the background to the master sequence. This is the same approximation usually used to simplify the analysis of the static single peaked landscape. A simplified one-dimensional version of the rate equation of the following form:

$$\dot{x}_0(t) = Q\sigma(t)x_0(t) - e(t)x_0(t), \tag{8}$$

where $Q = q^\nu$ is the copying fidelity of the whole genome and $e(t) = (\sigma(t) - 1)x_0(t) + 1$.

Figure 2 compares the concentration of the master sequence calculated by solving approximation 8 and by numerically integrating the full rate equation (2). The figure shows that the approximation is quite accurate.

Since this equation is one dimensional there is no time-ordering problem and it can be solved analytically for non-periodic peak fluctuations. Equation (8) can be transformed to a linear form by introducing a new variable $y(t) = [Q - x_0(t)]/[(1-Q)x_0(t)]$. This gives

$$\dot{y}(t) = 1 - [Q\sigma(t) - 1]y(t), \tag{9}$$

which can be solved. Substituting back gives the concentration of the master sequence

$$x_0(t) = x_0 \frac{\exp\left\{ \int_0^t [Q\sigma(s) - 1] dn \right\}}{1 + x_0 \int_0^t \exp\left\{ \int_0^s (Q\sigma(u) - 1) du \right\} [\sigma(s) - 1] ds}. \tag{10}$$

Since we are only interested in the long time behavior of the system, we can ignore transients carrying memory from initial values. Assuming $\exp \int_u^t [Q\sigma(s) - 1] ds \gg 1$ gives

$$x_0(t) = \frac{Q}{1 + (1-Q) \int_0^t \exp\left\{ \int_s^l [Q_\sigma(u) - 1] du \right\} ds}. \tag{11}$$

This is a generalization of the static expression for the asymptotic concentration:

$$x_0^s = \frac{Q\sigma - 1}{\sigma - 1}. \tag{12}$$

On a static single peaked fitness landscape there is a phase transition in the concentration distribution when the copying fidelity decreases below a critical value [2,13]. At high mutation rate the selective advantage of the master sequence due to its superior fitness is no longer strong enough for the gene sequences to be localized in sequence space. Instead they diffuse over the entire sequence space, and the distribution becomes approximately uniform. This is generally referred to as the error catastrophe or error threshold and is one of the main implications of the original quasispecies model. By making the same approximation as above, i.e., assuming no back mutations onto the master sequence, the static landscape error threshold can be shown to occur when $Q = 1/\sigma$. In other words, the transition occurs when the selective advantage of the master sequence no longer is able to compensate for the loss of offspring due to mutations. This can also be seen from Eq. (12) that defines the stationary distribution of the master sequence in the static case.

One has to be careful when discussing the error threshold on a fluctuating peak. The fitness can, for example, slowly move from being strong enough to localize the population around the peak to being so weak that the population delocalizes, and then back again. If we, however, consider an average over a time scale much longer than the fluctuation time of the fitness peak, a sensible definition of the error threshold can be made based on the average concentration of the master sequence. The time average of the concentrations can be found by rewriting Eq. (8) as differentials

$$\int \frac{dx_0}{x_0} = \int \{Q\sigma(t) - 1 - [\sigma(t) - 1]x_0(t)\} dt. \tag{13}$$

The concentration of the master sequence is positive. The left-hand side of Eq. (13) is, therefore, positive and the last term in the integral $-[\sigma(t) - 1]x_0(t)$ on the left-hand side is

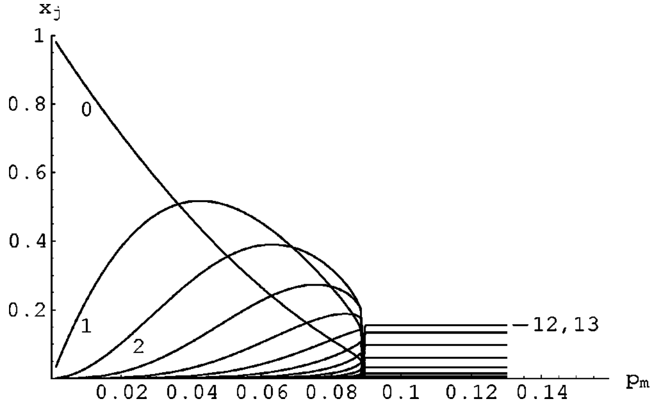

FIG. 3. Time-averaged quasispecies distribution is shown as a function of the error rate $p_m = q - 1$. The figure shows the numerical solution to the full rate equations. The fitness peak is defined as $\sigma(t) = 10 + 5\sin(t)$ and the genome length is $\nu = 25$. The error threshold is located at $p_m \approx 0.085$, corresponding to $Q_{\text{fc}} = 0.109$, which can be compared to the approximate value $Q_{\text{fc}} = 0.1$ predicted by Eq. (14).

negative. This implies that for $x_0(t)$ to be positive as time goes to infinity, we must assume $\int [Q\sigma(t) - 1]dt > 0$. The fluctuating time-dependent equivalent to the static error threshold is, therefore, given by

$$Q_{\text{fluct. crit.}} = \frac{1}{\langle \sigma \rangle_t}. \tag{14}$$

This shows that the error threshold on a fluctuating fitness peak is determined by the time average of the fitness, if the fluctuations are fast compared to the response time of the population.

Equation (11) indicates that the response time of the system is approximately given by $[Q\sigma(t) - 1]^{-1}$, i.e., the relative growth of the master sequence compared to the background. For the time average mentioned above to be an interesting parameter the fluctuations of the fitness peak must, therefore, be faster than this response time; only for this kind of environmental dynamics is it sensible to talk in terms of the average concentration of the master sequence. Thus, if the fluctuations occur on a time scale faster than the response time of the quasispecies, then the error threshold is defined by Eq. (14). For extremely slow changes the system will effectively be in equilibrium around the current value of the fitness. For slightly faster changes the response of the population will lag somewhat behind the changes in selective environment. In these cases it is more interesting to study the minimal concentration of the master sequence, which occurs when the fitness peak has a minimum (as we shall see later the phase shift decreases when the fluctuation frequency decreases).

When the full replicator equations for a rapidly fluctuating peak are numerically integrated, the time-averaged quasispecies distribution displays an error catastrophe at high error rates $\mu = 1 - q$. In Fig. 3 the fitness peak fluctuates periodically with $\sigma(t) = 10 + 5\sin(t)$. The average fitness is given by $\langle \sigma \rangle_t = 10$ and the genome length $\nu = 25$ and thus Eq. (14) predicts the error threshold to occur at $\mu = 0.088$, which agrees with the value found by numerically integrating the equations of motion directly. The analysis in this section demonstrates that by making the error tail approximation and reducing the dynamics to one-dimensional form, an analytic form exists for the error threshold on fast moving landscapes. This one-dimensional formulation removes the need to time order the changes in selective advantage of the landscape. This allows the integrals for the time history of the master-sequence concentration to be solved explicitly in Eq. (10).

IV. PHASE SHIFTS ON PERIODIC LANDSCAPES

To study how the master sequence responds to changes in the height of the fitness peak it is convenient to assume that the fluctuations are periodic. In this case it follows immediately from *Floquet's theorem* (also known as the *Bloch theorem* in statistical physics) that the solutions to the rate equations also are periodic with the same frequency. It then makes sense to speak of the amplitude of the oscillations in concentration and of the phase shift between the concentration and the fitness. It is intuitively clear that when the fitness peak is oscillating slowly (compared to the response time $[Q\sigma(t) - 1]^{-1}$) there will be a very small phase shift; the population will have time to reach an equilibrium about every value of $\sigma(t)$. The amplitude of changes in the master-sequence concentration will for the same reason, be as large as possible. This result, together with the time-averaging effect found in the previous section, indicates that the population responds to the driving of the environment with a low-pass filter effect. In one-dimensional population genetic models this phenomenon has been noted for some time [14–16,18]. Wilke, Ronnewinkal, and Martinetz [11] demonstrated via simulations that the same filtering occurred to quasispecies evolution on a periodically fluctuating single peak. Noting that the maxima and minima in concentration occurs when $\dot{x}_0 = 0$, we can find a relation between the phase shift (between the concentration and fitness fluctuations), and the amplitude of the fitness fluctuations. Let $t_{x\,\max}$ be the time when the concentration has a maximum. Similarly, the fitness is at a maximum at time $t_{\sigma\,\max}$. Thus the phase shift between the two is $\delta = t_{x\,\max} - t_{\sigma\,\max}$. From Eq. (8) the condition for the maximum value of x_0 during a full cycle can be derived

$$\max[x_0(t)] = \frac{Q\sigma(t_{\sigma\,\max} + \delta) - 1}{\sigma(t_{\sigma\,\max} + t_\delta) - 1}. \tag{15}$$

In general, there is no closed analytic expression for this phase shift (δ), or the response amplitude of the master-sequence concentration. When the fluctuations of the fitness peak is a small harmonic oscillation Eq. (15) becomes analytically tractable. For such fluctuations

$$\dot{y}(t) = 1 - [Q\sigma(t) - 1]y(t), \tag{16}$$

$$\sigma(t) = \overline{\sigma} + \epsilon \sin(\omega t). \tag{17}$$

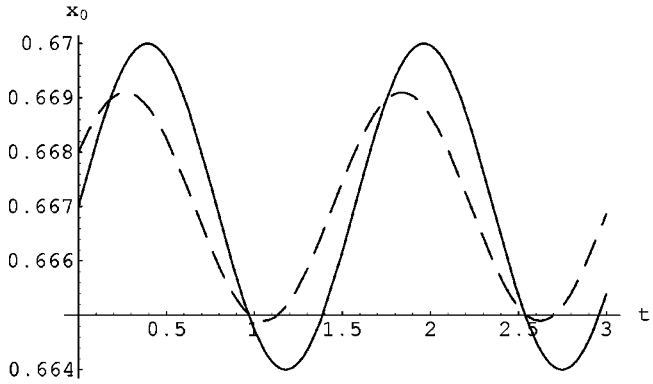

FIG. 4. Response in concentration of the master sequence (solid line) as the fitness peak oscillates according to $\sigma(t)=10+\sin(4t)$. The genomic copying fidelity is $Q=0.7$. The dashed line shows $\sigma(t)$, scaled to fit in the plot. Note the phase shift between the fitness function and the concentration response.

From Eq. (11) it is reasonable to assume the solution to be of the form $y(t)=1/(Q\overline{\sigma}-1)+u(t)$, where $u(t)$ is small compared to the average. Ignoring higher-order terms Eq. (8) can be written in terms of the perturbation $u(t)$ as

$$\dot{u}(t)=(1-Q\overline{\sigma})u(t)-\frac{\epsilon Q\sin(\omega t)}{Q\overline{\sigma}-1}. \tag{18}$$

This differential equation can be solved to obtain

$$u(t)=-\frac{\epsilon Q}{(Q\overline{\sigma}-1)\sqrt{(Q\overline{\sigma}-1)^2+\omega^2}}\sin(\omega t-\delta), \tag{19}$$

$$\tan(\delta)=\frac{\omega}{Q\overline{\sigma}-1}. \tag{20}$$

In Eq. (19) and (20) transients have been ignored since they decay exponentially as $e^{-(Q\overline{\sigma}-1)t}$. Thus the frequency of the oscillations is normalized by the (average) response rate of the population $Q\overline{\sigma}-1$.

Substituting this back into the expression for $x_0(t)$ gives

$$x_0(t)=\frac{\overline{x}}{1-\frac{\epsilon(1-Q)\sin(\omega t-\delta)}{(\overline{\sigma}-1)\sqrt{(Q\overline{\sigma}-1)^2+\omega^2}}}, \tag{21}$$

where $\overline{x}=(Q\overline{\sigma}-1)/(\overline{\sigma}-1)$.

The characteristic behavior of a low-pass filter is clearly shown in Eq. (20) and (21). As the frequency of the fluctuations increases, the amplitude of the concentration response decreases and the phase shift converges to $\pi/2$. Figure 4 shows how a population responds to harmonic oscillations of the fitness peak. The phase shift makes the concentration of the master sequence reach its maximum when the actual fitness has already decreased below maximum.

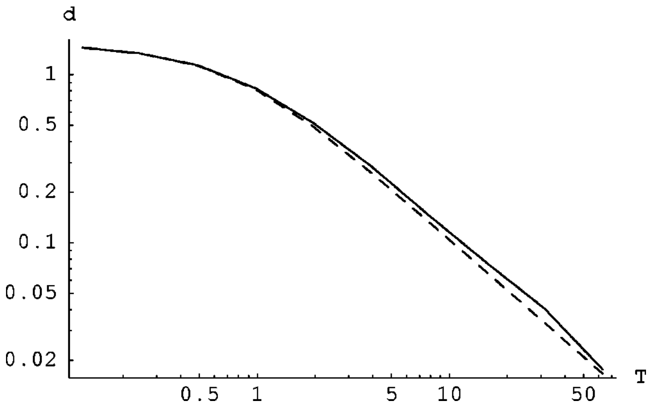

FIG. 5. Phase shift as a function of the period $T=2\pi/\omega$. The dashed line is a prediction using Eq. (20) and the solid is derived by numerically solving the rate equations (2). Parameters used are $\sigma=10+\sin(\omega t)$, $Q=0.7$, and $\nu=25$.

V. CONCLUSIONS

In this paper, we have shown that the time dynamics of a quasispecies on a fluctuating peak can be studied under the standard no back-mutation approximation. The general time-ordering problem stemming from a time-dependent landscape disappears since the rate equation becomes one dimensional. We show that the time-dependent equivalent to the static error threshold is determined by the time average of the fluctuations of the fitness peak. An expression for the typical response time for a population is given in terms of copying fidelity and selection pressure. We also show that for small periodic fluctuations the time dynamics of the population has a phase shift and a low-pass filter amplitude response. Analytic expressions for the phase shift and the amplitude are derived in the special case of small harmonically oscillating fluctuations.

ACKNOWLEDGMENTS

When doing this work N.S. and M.N. were supported by SFI core funding grants. N.S. would also like to acknowledge the support of Mats Nordahl at Chalmers University of Technology while preparing this manuscript. We also thank Mats Nordahl for valuable comments and discussions.

---

[1] M. Eigen, Naturwissenschaften **58**, 465 (1971).
[2] M. Eigen and P. Schuster, Naturwissenschaften **64**, 541 (1977).
[3] P. Schuster, Physica D **16**, 100 (1986).
[4] P. Schuster and K. Sigmund, Ber. Bunsenges. Phys. Chem. **89**, 668 (1985).
[5] I. Leuthäusser, J. Chem. Phys. **84**, 1884 (1986).
[6] P. Tarazona, Phys. Rev. A **45**, 6038 (1992).
[7] J. Swetina and P. Schuster, Bull. Math. Biol. **50**, 635 (1988).
[8] M. Nowak and P. Schuster, J. Theor. Biol. **137**, 375 (1989).
[9] M. Eigen and P. Schuster, *The Hypercycle—A Principle of Natural Self-Organization* (Springer, Berlin, 1979).
[10] M. Eigen, J. McCaskill, and P. Schuster, Adv. Chem. Phys. **75**,

149 (1989).

[11] C. O. Wilke, C. Ronnewinkel, and T. Martinetz, in *Proceedings of ECAL'99*, edited by H. Lund and R. Kortmann Lecture Notes in Computer Science (Springer-Verlag, Heidelberg, 1999); p. 417 e-print physics/9904028.

[12] C. O. Wilke, C. Ronnewinkel, and T. Martinetz, Phys. Rep. **349**, 395 (2001).

[13] J. Maynard Smith and E. Szathmáry, *The Major Transitions in Evolution* (W. H. Freeman, Oxford, 1995).

[14] K. Ishii, H. Matsuda, Y. Iwasa, and A. Saskai, Genetics **121**, 163 (1989).

[15] B. Charlesworth, J. Hered **84**, 345 (1993).

[16] R. Lande and S. Shannon, Evolution (Lawrence, Kans.) **50**, 434 (1996).

[17] M. Nilsson and N. Snoad, Phys. Rev. Lett. **84**, 191 (2000).

[18] A. J. Hirst, in *Proceedings of the Fourth European Conference on Artificial Life* (ECAL97), edited by P. Husbands and I. Harvey (MIT/Bradford Books, Cambridge, MA, 1997), pp. 425–431.