Detailed analysis of an Eigen quasispecies model in a periodically moving sharp-peak landscape

Armando G. M. Neves*
Depto. de Matemática, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil
(Received 6 May 2010; published 24 September 2010)

The Eigen quasispecies model in a periodically moving sharp-peak landscape considered in previous seminal works [M. Nilsson and N. Snoad, Phys. Rev. Lett. **84**, 191 (2000)] and [C. Ronnewinkel *et al.*, in *Theoretical Aspects of Evolutionary Computing*, edited by L. Kallel, B. Naudts, and A. Rogers (Springer-Verlag, Heidelberg, 2001)] is analyzed in greater detail. We show here, through a more rigorous analysis, that results in those papers are qualitatively correct. In particular, we obtain a phase diagram for the existence of a quasispecies with the same shape as in the above cited paper by C. Ronnewinkel *et al.*, with upper and lower thresholds for the mutation rate between which a quasispecies may survive. A difference is that the upper value is larger and the lower value is smaller than the previously reported ones, so that the range for quasispecies existence is always larger than thought before. The quantitative information provided might also be important in understanding genetic variability in virus populations and has possible applications in antiviral therapies. The results in the quoted papers were obtained by studying the populations only at some few genomes. As we will show, this amounts to diagonalizing a 3×3 matrix. Our work is based instead in a different division of the population allowing a finer control of the populations at various relevant genetic sequences. The existence of a quasispecies will be related to Perron-Frobenius eigenvalues. Although huge matrices of sizes 2^{ℓ}, where ℓ is the genome length, may seem necessary at a first look, we show that such large sizes are not necessary and easily obtain numerical and analytical results for their eigenvalues.

DOI: 10.1103/PhysRevE.82.031915 PACS number(s): 87.23.Kg, 87.10.Ca

I. INTRODUCTION

A *quasispecies* is, in rough terms, an equilibrium distribution of similar self-replicating entities. The term was introduced in the early 1970s by Eigen [1] in a context of prebiotic evolution, in which entities were identified, e.g., with RNA molecules. Later on the same kind of models came to be thought as an interesting framework for understanding the evolution of viruses, both RNA and DNA based [2]. Besides Eigen's model, a different quasispecies model was formulated by Crow and Kimura [3], also known as ParaMuSe model. Although the two classes of models are technically different, results are similar (see, e.g., [4,5]). There is nowadays a vast literature on quasispecies models with many contributions given by physicists. The first works dealt mostly with static fitness landscapes. For this part of the literature, we refer the reader to reviews [6–8] where relevant and primary references are acknowledged. Remarkable also are some works in which quasispecies models are identified with Ising spin systems [9,10] or quantum spin chains [11–14]. Such identifications led not only to new insights, but also to approximate and exact analytical results.

The work on dynamic landscapes began by the year 2000 [15–21]. As arbitrary time dependences are difficult to work with, almost all results are on some kind of periodic dependence. Even restricting to periodic dynamics, there are two typical situations. One is when fitnesses depend periodically on time, studied, e.g., in [17–20]. The other is when the whole landscape moves through genome space periodically in time, considered in [21,15,16] and in the present work.

More concretely, in this paper we are going to analyze a dynamic version of the sharp-peak fitness landscape (SPL), in which the peak moves periodically in genome space to a random nearest neighbor, i.e., Hamming distance (HD) equal to 1, of the previously occupied position. This is almost exactly the same situation studied in the remarkable papers by Nilsson and Snoad [15] and Ronnewinkel *et al.* [21]. The only difference with respect to [15] is that we are going to consider discrete time, whereas they use continuous time. This is an inessential difference resulting only from a matter of taste, and we will present when necessary the results of [15] in a discrete-time version.

In [21], the SPL is called needle in the haystack; discrete time is used as here, but authors are more concerned with deterministic motion of the master sequence. Nonetheless, they comment briefly on the stochastic motion of the master sequence, concluding that "the overall behavior is similar."

Despite some differences, both [15,21] arrive at the same results, which we summarize in Sec. II, using very similar assumptions. We should note however that the slight difference between deterministic and stochastic motions of the master sequence, a subject to be commented on at Sec. III, was brought to our attention only by [21].

The motivation behind a moving fitness landscape is that viruses need to mutate in order to survive the attack of their hosts immune systems, which is the moving environment in which they live. As hosts acquire immunity against strains of a virus, an optimal virus genome at a certain time will no longer be optimal when most host individuals will have acquired immunity against viruses carrying that genome.

Before we go on, we should point that whereas we consider the immune system as an external influence on the virus population, there exist different ways of describing the interaction of viruses and immune systems, e.g., [22–24].

We also acknowledge recent work by Ancliff and Park on periodically moving fitness landscapes [25–27]. In [26,27], they dealt with a model very different from the one we treat

*aneves@mat.ufmg.br

here, in which the master sequence alternates periodically between two sequences very far apart in HD. On the other hand, in [25] they used results from [13] based on the identification with a quantum spin chain [11] to obtain maximum, minimum, and optimal mutation rates for the existence of a quasispecies in a ParaMuSe model in which the peak in a SPL moves periodically to a sequence apart a certain number of HD units, with this number being not necessarily equal to 1. Although their results have some resemblance to ours, the techniques are very different, and the models cannot be thought of as equivalent. In our opinion, the major difference with respect to our results is that in their model, sequences which are not master are taken as having null fitness. In particular, the region of low selective advantage of the master sequence with respect to the other sequences, which we also treat, is completely excluded from their analysis.

The analysis to be performed here differs from the ones in [15,21] in that we do not divide the population into error classes and do not rely on certain approximations they did. In order not to have to deal with the population at all sequences in the huge genome space, we choose to track only the population in all genetic sequences which are going to be master sequence at some time. This amounts already to a substantial increase in complexity with respect to [15] or [21], but we are able to obtain very good numerical results without using overwhelmingly large matrices, and approximate formulas are also available. A first result is a mathematically sounder confirmation of the existence of upper and lower thresholds for the mutation rate between which a quasispecies exists. More importantly, we show in our analysis that the existence of a quasispecies is possible in a wider region of the space of parameters than found in [15] or [21]. The quantitative difference between the present results and those of [15] or [21] is important for small values of the selective advantage of the master sequence with respect to the other sequences.

It has been suggested, see [28] and references therein, that the error catastrophe present in quasispecies models could be useful as an antiviral strategy. The idea is that instead of fighting viruses with conventional drugs, a different possibility is to induce quasispecies extinction by using mutagen drugs to increase viruses' mutation rate beyond the upper threshold. The necessity of fully understanding how mutagens act justifies our quest for a good quantitative knowledge of error thresholds. We also believe that the more detailed information we provide on populations at many genetic sequences besides the master sequence could be useful when trying to understand genetic variability in real virus populations.

The paper is organized as follows. In Sec. II we sketch the basic notions on Eigen quasispecies models, the SPL, and provide a statement of the discrete-time version of the analyses by Nilsson and Snoad [15] and Ronnewinkel *et al.* [21]. In Sec. III we state the fundamentals of our own finer analysis of the same model. In Sec. IV we present numerical and approximate analytical results which show the differences between our results and those of [15,21]. Finally, in Sec. V we summarize and discuss our findings.

II. EIGEN'S MODEL AND THE ANALYSES OF NILSSON AND SNOAD AND RONNEWINKEL ET AL.

Consider a population of viruses. Each individual in this population is characterized by a sequence $\sigma=(s_1,s_2,\ldots,s_\ell)$ of ℓ letters chosen from an alphabet with N symbols. The sequence should be thought as the individual's genome. Individuals reproduce asexually at a rate dependent on their genome σ and in doing so produce new individuals generally with the same genome, but mutations are allowed and constitute a basic ingredient of the model.

Let Λ denote the set of all N^ℓ possible in principle genomes. In nature, $N=4$, as the letters of the alphabet are the bases A, C, G, and T for DNA-based organisms or A, C, G, and U for RNA-based ones. In the physics literature a common practice is to simplify things and use a binary alphabet. On the other hand, even in an individual as simple as a virus we have $\ell\sim 10^3-10^5$ and much larger values for more complex organisms [16]. We will assume throughout that $\ell\gg 1$, so that Λ is a huge set, much larger than any conceivable population.

Although this will be given a quantitative definition in a while, we will say that a population is a quasispecies if it does not spread too much in the genome space Λ, i.e., if it consists of individuals with a *wild-type* or *master sequence* genome σ_0 along with mutant individuals with genomes not very distant from the master sequence. The relevant concept of distance to be used here is the *Hamming distance* $d(\sigma,\sigma')$, defined as the minimum number of letter substitutions to be performed in sequence σ to make it coincide with σ'.

If $p_\sigma(t)$ denotes the number of individuals in the population with genome σ at time t, then the basic equation of Eigen quasispecies model, which describes the replication dynamics, is

$$p_\sigma(t+1)=\sum_{\sigma'\in\Lambda} W_{\sigma,\sigma'}f(\sigma')p_{\sigma'}(t). \tag{1}$$

In this equation, we consider time as a discrete variable, with each generation living one unit of time, individuals producing offspring at the end of their lives, and being replaced by the new generation. $f(\sigma')$ is the *fitness* of an individual, i.e., the number of offspring produced by it, which we assume dependent only of its genome. $W_{\sigma,\sigma'}$ is the probability that an individual of type σ' has offspring of type σ due to mutations. In spite of using the term probability here, we are supposing that populations are so large that statistical fluctuations are negligible, and the model is completely deterministic. For an account of quasispecies theories for finite populations, see, e.g., [29].

An important ingredient of quasispecies models is the *fitness landscape*, i.e., specification of the function f in the above equation. Many static fitness landscapes such as sharp peak, Fujiyama, and random have been studied [6]. The *sharp-peak* landscape is defined by the relative fitness

$$f(\sigma)=\begin{cases}1+k, & \text{if} \quad \sigma=\sigma_0\\ 1, & \text{if} \quad \sigma\neq\sigma_0,\end{cases} \tag{2}$$

where $k>0$ is the *selective advantage* of the master sequence with respect to all other sequences.

Notice now that Eq. (1) is difficult to analyze without further simplifications, because it is a system of N^ℓ equations; the matrix of $W_{\sigma\sigma'}$ might be very complex and the fitness landscape f complicates things even more. A first simplification we introduce is to consider that each letter of the genome has a uniform probability of mutating to any other letter and that the genome length ℓ is always conserved. Thus, we have $W_{\sigma,\sigma'}=\mu^{d(\sigma,\sigma')}(1-\mu)^{\ell-d(\sigma,\sigma')}$. As the *mutation rate per base and per replication* μ is very small in real-life organisms, of order 10^{-7}–10^{-11} [16], a further simplification is to neglect terms $O(\mu^n)$ with $n\geq 2$ and take

$$W_{\sigma\sigma'}=\begin{cases}1-\ell\mu, & \text{if} \quad d(\sigma,\sigma')=0\\ \mu, & \text{if} \quad d(\sigma,\sigma')=1\\ 0, & \text{if} \quad d(\sigma,\sigma')>1.\end{cases} \tag{3}$$

These simplifications are also adopted in [15,21].

A further simplification also adopted in [15,21], but *not* in this paper, is to divide the set Λ into *error classes*. Instead of studying the number of individuals in all N^ℓ genomes, it is easier to study only the number of individuals in the much smaller set of error classes. The ith error class $\Gamma_i(t)$ is defined as the set of all genomes in Λ with HD from the master sequence equal to i $(i=0,1,2,\ldots,\ell)$. We are anticipating here that as the master sequence genome σ_0 may depend on t, so do the error classes.

Another simplification typical of the SPL and also used in [15,21] is that of *no backmutations*. The number of elements in Γ_{i+1} is $\frac{\ell-i}{i+1}$ times the number of elements in Γ_i. Then, as $\ell\gg 1$, a new mutation in an individual at Γ_i with small i will most probably result in an individual in Γ_{i+1} and not in Γ_{i-1}. For i close to ℓ, instead, we expect the contrary. But, as error classes with very large i are expected to be scarcely populated, the no-backmutation simplification adopted in [15,21] and many other papers is that mutations will always drive individuals to larger order error classes.

The situation studied in [15,21] is the one in which the static SPL (2) holds for a *cycle* of τ units of time (generations) and is followed by an instantaneous *shift* of the master sequence to another sequence in genome space at HD equal to 1. Then for the next cycle of τ generations we use again Eq. (2) with respect to the new master sequence, and then another shift follows and so on.

Let then $p_{\mathrm{NS}}(t)=(A(t),B(t),C(t))$, where $A(t)$ is the number of individuals in error class $\Gamma_0(t)$, i.e., in the master sequence at time t, $B(t)$ is the number of individuals in the specific genome in $\Gamma_1(t)$ which will become master sequence after the next genome shift and $C(t)$ is the number of individuals in all other sequences. By using the simplified mutation matrix (3), the no-backmutation simplification, and Eq. (2), we find that for any instants of time between successive shifts of the master sequence we have

$$p_{\mathrm{NS}}(t+1)=E_{\mathrm{NS}}p_{\mathrm{NS}}(t), \tag{4}$$

where

$$E_{\mathrm{NS}}=\begin{pmatrix}(1-\beta)(1+k) & 0 & 0\\ \frac{\beta}{\ell}(1+k) & 1-\beta & 0\\ \frac{\beta(\ell-1)}{\ell}(1+k) & \beta & 1\end{pmatrix}, \tag{5}$$

and we introduced the *genome mutation rate* $\beta=\mu\ell$. Of course, we may calculate population evolution before the first shift by $p_{\mathrm{NS}}(\tau_-)\equiv E_{\mathrm{NS}}^{\tau}p_{\mathrm{NS}}(0)$.

The effect of the master sequence shift, as in the analyses of [15,21], is to instantly replace the population vector at time τ_- by taking $A(\tau_+)=B(\tau_-)$, because sequence of population B will be master for the next τ generations, and

$$B(\tau_+)=0, \tag{6}$$

because the sequence which will become master at $t=2\tau$ is most probably a sequence which was in $\Gamma_2(t)$ for $0\leq t<\tau$, and it is assumed that this error class is scarcely populated. In matrix language,

$$p_{\mathrm{NS}}(\tau_+)\equiv S_{\mathrm{NS}}p_{\mathrm{NS}}(\tau_-)\ , \tag{7}$$

where

$$S_{\mathrm{NS}}=\begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 0\\ 1 & \ell-1 & 1\end{pmatrix}. \tag{8}$$

We take then $p_{\mathrm{NS}}(\tau_+)$ as initial condition, use again Eq. (4) for the next cycle, and so on. The above equations define what we will call *Nilsson-Snoad (NS) dynamics*.

Another assumption of [15,21] in Eq. (8) is that all ℓ sequences in Γ_1 are equally populated, and hence $B(t)$ is just a representative of any sequence in Γ_1. The total population is then given by $N(t)=A(t)+\ell B(t)+C(t)$.

The *frequencies* (or concentrations) are defined as $a(t)=A(t)/N(t)$, $b(t)=B(t)/N(t)$, and $c(t)=C(t)/N(t)$. It can be seen that the frequency vector converges under the NS dynamics either to an equilibrium point or to an orbit of period τ. For some values of the parameters β,k,τ,ℓ, the limiting frequencies are the trivial equilibrium point (0,0,1), and for the other values the nontrivial orbit starts just after the shift of the master sequence with a vector in the form $(a_*,0,1-a_*)$ with $0<a_*<1$. In the former situation, all individuals in the population are finally scattered to the higher error classes, and genetic structure of the population is lost. In the latter, some genetic structure is preserved to infinite time. In this case we will say that a quasispecies exists because the long-term population consists of individuals in the master sequence along with mutants in the other classes providing the genetic variability necessary for adaptability.

The condition which separates the above cited situations is the value of the largest eigenvalue of matrix $S_{\mathrm{NS}}E_{\mathrm{NS}}^{\tau}$. It is easy to see that 0 and 1 are always eigenvalues of this matrix, with the latter having (0,0,1) as eigenvector. The third eigenvalue is

$$\lambda_{\mathrm{NS}} = \frac{[(1+k)^{\tau}-1](1+k)}{k\ell}\beta(1-\beta)^{\tau-1}, \tag{9}$$

which may be larger or smaller than 1. The condition for the existence of a quasispecies is then

$$\lambda_{\mathrm{NS}} > 1, \tag{10}$$

so that the dynamics is dominated by the nontrivial eigenvector related to the largest eigenvalue λ_{NS}. Equation (10) is to be taken as the discrete-time analog of inequality (6) of [15]. With slight differences and already in a discrete-time version, the above inequality is present also in [21].

In particular, as shown in [15,21], if k, τ, and ℓ are fixed, with k and τ being large enough and ℓ small enough, then there exist thresholds $0<\beta_l<\beta_u$ such that a quasispecies exists in the range $\beta \in (\beta_l, \beta_u)$. For other values of k, τ, and ℓ a quasispecies will not exist because λ_{NS} will be smaller than 1 for all β's. A static version of the upper threshold β_u was already known for the static SPL (see, e.g., [6]), but the lower threshold β_l is a novelty of the dynamic case. If $\beta > \beta_u$, mutations are too frequent, and natural selection is not able to overcome them. If $\beta<\beta_l$, mutations are so rare that population is not able to follow the periodic shift of the master sequence in genome space. In the first case, we talk of extinction due to the *error catastrophe* and in the second of *adaptability catastrophe* [22].

III. FINER ANALYSIS OF THE DYNAMIC SPL

The main point of the present paper is that we are able to analyze the same situation as in [15,21] with less assumptions or approximations and providing a more detailed account of the composition of the virus population. In particular, we think that taking Eq. (6) should approximately hold only when k is very large. Moreover, the motion of the master class breaks the symmetry among sequences in error classes for a very simple reason: for small i there exist in $\Gamma_i(t)$ one sequence which was master in previous cycles and sequences which have never been masters. Finally, another information lacking in the analysis in [15,21] is about populations in error classes Γ_i, with $i \geq 2$. Such information is certainly useful when trying to assess genetic variability of virus populations.

The basic idea of our analysis is that of keeping track at all times of the population in all sequences which will be master at some time. In particular, we do not assume that all sequences in $\Gamma_1(t)$ are equally populated and that the population of sequences in $\Gamma_2(t)$ is negligible.

If we think of the genome space Λ as the vertices of an ℓ-dimensional cube with large ℓ, then the successive positions occupied by the master sequence are a random walk in which at every τ generation a step is made to a nearest-neighbor vertex chosen with uniform probability. If $\sigma_1, \sigma_2, \ldots$ are the successive positions of the master sequence, then with a probability of 1 the master sequence will return some time to σ_1, but the number of steps for that is of order 2^{ℓ}. More exactly, there is a small probability equal to $1/\ell$ that the master sequence returns to σ_1 after only two steps. In case such an event occurs, we let the walk proceed and wait until the next return of the master sequence to σ_1, probably after a number of steps of order 2^{ℓ}. In fact, the probability of the master sequence returning to σ_1 at the fourth step is $2/\ell^2(1-1/\ell)$, much smaller, and so on. Anyway, in average once each ℓ steps, the master sequence will perform a two-step backjump to a previous position. As $2^{\ell} \gg \ell$, we may expect many two-step backjumps in the time of order 2^{ℓ} until the master sequence returns to σ_1. Backjumps in four or more steps will also occur many times at time scales as long as 2^{ℓ}, but they are much less frequent.

Having said that, let $\sigma_1, \sigma_2, \ldots, \sigma_M, \sigma_{M+1}$ denote the successive sequences in Λ which will be master at some time with the periodic boundary condition $\sigma_{M+1} \equiv \sigma_1$, so that $d(\sigma_j, \sigma_{j+1})=1$, for $j=1,2,\ldots,M$. By the above arguments, M is a random variable with typical values of order 2^{ℓ}. As ℓ is large, then in typical realizations of the random walk of the master sequence, we will have $d(\sigma_i, \sigma_j) \geq 2$ for most pairs i,j with $|i-j| \geq 2$. An obvious exception is when $i=1, j=M$ or vice versa and when backjumps occur. Ronnewinkel *et al.* in [21] worked most of the time with what they called *regular motions* of the master sequence, in which backjumps do not occur by definition. They also observed that backjumps were present in their simulations of stochastic motion of the master class, but these were not able to change the long-time behavior of the model. We will follow the same road and suppose for the time being that the motion of the master sequence is random, but without backjumps. In Sec. V we will justify why including backjumps will not modify the long-time results.

Since we are going to analyze population at M sequences in Λ, it follows that the method of analysis we are now introducing is located at an intermediate point between analyzing populations at all sequences in Λ, which would be unnecessarily detailed, and studying only populations in error classes Γ_0 and Γ_1, as in [15,21].

Let $p_i(t)$ denote the number of individuals with genome σ_i at time t, $p(t)$ be the vector $(p_1(t), p_2(t), \ldots, p_M(t))$ in $\mathbb{R}^M$, and $q(t)$ be the total number of individuals in all sequences in Λ which will not be masters at any time. Let also $\hat{p}(t)$ be the vector in $\mathbb{R}^{M+1}$ defined as $\hat{p}(t)=(p(t), q(t))$. For the sake of further reference, we will call vector $p(t)$ the *relevant population* and $\hat{p}(t)$ the *total population*. The extra component $q(t)$ of $\hat{p}(t)$ will be called the *background population*.

For the first τ generations, σ_1 will be the master sequence. During this time, i.e., for $0 \leq t \leq \tau-1$, Eigen's equation (1) with the SPL (2) and mutation matrix (3) can be written in matrix form as

$$\hat{p}(t+1) = \hat{E}_1 \hat{p}(t), \tag{11}$$

where $\hat{E}_1$ is the $(M+1)\times(M+1)$ matrix given in block form by

$$\hat{E}_1 = \begin{pmatrix} & & & 0 \\ & E_1 & & 0 \\ & & & 0 \\ & & & \vdots \\ & & & 0 \\ \left(1-\frac{2}{\ell}\right)\beta(1+k) & \left(1-\frac{2}{\ell}\right)\beta & \cdots & \left(1-\frac{2}{\ell}\right)\beta \quad 1 \end{pmatrix}. \tag{12}$$

In the above equation, E_1 is the $M \times M$ matrix for evolution only of the relevant population, i.e., $p(t+1)=E_1 p(t)$, and is explicitly given by

$$E_1 = \begin{pmatrix} (1-\beta)(1+k) & \frac{\beta}{\ell} & 0 & 0 & \cdots & 0 & \frac{\beta}{\ell} \\ \frac{\beta}{\ell}(1+k) & 1-\beta & \frac{\beta}{\ell} & 0 & \cdots & 0 & 0 \\ 0 & \frac{\beta}{\ell} & 1-\beta & \frac{\beta}{\ell} & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\beta}{\ell}(1+k) & 0 & 0 & 0 & \cdots & \frac{\beta}{\ell} & 1-\beta \end{pmatrix}. \tag{13}$$

More generally, for $(j-1)\tau \le t \le j\tau - 1$ we will have

$$\hat{p}(t+1) = \hat{E}_j \hat{p}(t), \tag{14}$$

with

$$\hat{E}_j = \hat{S}^{j-1} \hat{E}_1 (\hat{S}^{-1})^{j-1}. \tag{15}$$

The $(M+1)\times(M+1)$ matrix $\hat{S}$ implements the master sequence shift by shifting 1 unit to the right, taking into account the periodic boundary condition, the first M components of the total population $\hat{p}$, and leaving fixed the last component, i.e.,

$$\hat{S} = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 & 0 \\ 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 \end{pmatrix}. \tag{16}$$

As we will need it in a while, we also define an $M \times M$ matrix S which is simply $\hat{S}$ with the last row and the last column deleted, i.e., the restriction of $\hat{S}$ to the relevant population.

The above Eqs. (14) and (15) define our version of the discrete-time Eigen quasispecies model with the SPL and a periodical shift of the master sequence. At a time in which σ_i is the master sequence, both p_{i-1} and p_{i+1} are populations in the first error class, but they are not treated collectively as in [15,21]. Notice that offspring of individuals in the relevant population may be in the background due to mutation, but for simplicity we do not allow offspring of individuals in q to be relevant. This is indeed a sort of the no-backmutation approximation, but milder than the usual form in [15,21]. In fact we account partly for the possibility of backmutations, because individuals in σ_i are allowed to mutate to σ_{i-1}.

With the above definitions, it is easy to see that by the end of the first cycle of τ generations, i.e., by the end of the time in which σ_1 is the master sequence, the population vector is given by $\hat{p}(\tau)=\hat{E}_1^{\tau} p(0)$. It should be also clear that a steady state for the *frequencies* is a population $\hat{v}$ obeying

$$\hat{E}_1^{\tau} \hat{v} = \hat{\lambda} \hat{S} \hat{v}, \tag{17}$$

where $\hat{\lambda}$ is the growth factor for the whole population in the time of τ generations. In spite of $\hat{v}$ not being a steady-state population if $\hat{\lambda} \neq 1$, for ease of reference, we will still call it a *steady-state population*. By Eq. (17), the steady-state population $\hat{v}$ is an eigenvector of $\hat{S}^{-1}\hat{E}_1^{\tau}$, and $\hat{\lambda}$ is the corresponding eigenvalue.

Of course, identification of an eigenvector with a population will be possible only if $\hat{v}$ is a *non-negative vector*, in the sense of having all its components non-negative. As matrices E_1, $\hat{E}_1$, S, and $\hat{S}$ are also non-negative, then E_1^{τ}, $S^{-1}E_1^{\tau}$, $\hat{E}_1^{\tau}$, and $\hat{S}^{-1}\hat{E}_1^{\tau}$ are non-negative. We remind now the reader of the most important result on non-negative matrices, the Perron-Frobenius (PF) theorem (see, e.g., Theorem 4.2 in Chap. I of [30]). This theorem guarantees that for any non-negative matrix G there exists $r>0$ such that all eigenvalues $\lambda \in \mathbb{C}$ of G are such that $|\lambda| \le r$ and that r itself is an eigenvalue of G, which we will call the *PF eigenvalue*, and its corresponding eigenvector can be taken as non-negative.

Using the PF theorem and other related results, see again [30], we can prove some important facts:

(1) If λ_{PF} denotes the PF eigenvalue of $S^{-1}E_1^{\tau}$, then λ_{PF} is a simple root of the characteristic polynomial of $S^{-1}E_1^{\tau}$, and all other eigenvalues of this matrix have absolute values strictly smaller than λ_{PF}. Furthermore, we can associate with the eigenvalue λ_{PF} an eigenvector v_{PF} having strictly positive components, and positive multiples of v_{PF} are the only eigenvectors of $S^{-1}E_1^{\tau}$ with non-negative components. These facts depend on the irreducibility and primitivity of $S^{-1}E_1^{\tau}$.

(2) If $\hat{\lambda}_{\mathrm{PF}}$ denotes the PF eigenvalue of $\hat{S}^{-1}\hat{E}_1^{\tau}$, then

$$\hat{\lambda}_{\mathrm{PF}} = \max\{\lambda_{\mathrm{PF}}, 1\}. \tag{18}$$

(3) Despite the reducibility of $\hat{S}^{-1}\hat{E}_1^{\tau}$, we still have uniqueness of its PF eigenvector if $\lambda_{\mathrm{PF}} \neq 1$. If $\lambda_{\mathrm{PF}} < 1$, the only eigenvectors of $\hat{S}^{-1}\hat{E}_1^{\tau}$ with eigenvalue $\hat{\lambda}_{\mathrm{PF}}=1$ are the multiples of the trivial $\hat{v}_{\mathrm{PF}} \equiv (0,0,\ldots,0,1)$. If $\lambda_{\mathrm{PF}} > 1$, the only non-negative eigenvectors of $\hat{S}^{-1}\hat{E}_1^{\tau}$ are multiples of a positive vector $\hat{v}_{\mathrm{PF}}$ which coincides with v_{PF} in its relevant components. The last component of $\hat{v}_{\mathrm{PF}}$ in the case $\lambda_{\mathrm{PF}} > 1$ is

$$q = \frac{v_{\mathrm{PF}} \cdot w}{\lambda_{\mathrm{PF}} - 1}, \tag{19}$$

where w is the vector in $\mathbb{R}^M$ whose components are $w_j = (\hat{E}_1^{\tau})_{M+1,j}$.

We refer the reader to [31] for the mathematical details of the proofs of the following facts. From these facts it follows that for any initial non-negative population $\hat{p}(0)$ not a multiple of $(0,0,\ldots,0,1)$, $\hat{p}(j\tau)$ will converge as $j\to\infty$ to a trivial multiple of $(0,0,\ldots,0,1)$ if $\lambda_{\mathrm{PF}}<1$ or to some nontrivial vector having all components positive if $\lambda_{\mathrm{PF}}>1$. The border case $\lambda_{\mathrm{PF}}=1$ is still an open mathematical problem, but we do not think it is biologically relevant. In other words, in the long term any initial population not entirely concentrated on the background sequences will preserve some genetic structure if $\lambda_{\mathrm{PF}}>1$ and lose it if $\lambda_{\mathrm{PF}}<1$.

The condition for the existence of a quasispecies in the present analysis is thus $\lambda_{\mathrm{PF}}>1$. Equation (19) above shows that the background population q in the steady-state situation tends to become larger than the total relevant population when $\lambda_{\mathrm{PF}}\searrow 1$. We will see ahead that in our analysis, as in [15,21], for some values of k,τ,ℓ there exists a range (β_l,β_u) in β where a quasispecies exists. As λ_{PF} becomes equal to 1 at the boundaries β_l,β_u of this range, the above comment shows that the steady-state background population may be much larger than the steady-state population at the master sequence if one is in the quasispecies existence range, but close to its boundaries. It might seem odd to call such a population dominated by individuals with background genomes a quasispecies, but what characterizes a quasispecies is not the relative size of the master sequence and background populations, but rather the existence of a nonzero fraction of the population which maintains genetic structure in the long term.

In the regime of quasispecies existence, Eq. (17) and (18) show that $\lambda_{\mathrm{PF}}^{1/\tau}$ is the mean fitness of the population. We will obtain ahead Eq. (23), which can then be used to calculate the optimal mutation rate β_{opt} for fixed k, τ, and ℓ.

IV. SOME RESULTS

We have shown that dividing the population into the classes which are going to be master sequences at some time and a background also reduces the problem of studying the existence of a quasispecies in the moving SPL to a calculation of eigenvalues. That would be useless if we would be forced to use matrix sizes M of order 2^{ℓ}, as may seem necessary at a first look.

In Fig. 1 we plot the numerically calculated PF eigenvalue of $S^{-1}E_1^{\tau}$ for small values of M. The results illustrate the existence of upper and lower thresholds for the mutation rate which allow the existence of a quasispecies. We should point out at this moment that the numerical results also show that for values of β starting from zero up to a value $\bar{\beta}$ we have $\lambda_{\mathrm{PF}}>1$, with the value of $\bar{\beta}$ depending heavily on M, as shown by the picture. We will argue later that $\bar{\beta}\to 0$ in the limit $M\to\infty$, showing that the existence of this initial interval is an artifact due to the smallness of M and should not be thought of as a new phenomenon. On the contrary, for larger values of β the numerical calculation gives values for λ_{PF} rather independent of M with good coincidence of the values for $M=20$ and $M=100$. We have also calculated values for $M=200$, but they are indistinguishable from those for $M=100$ and were not included in the plot.

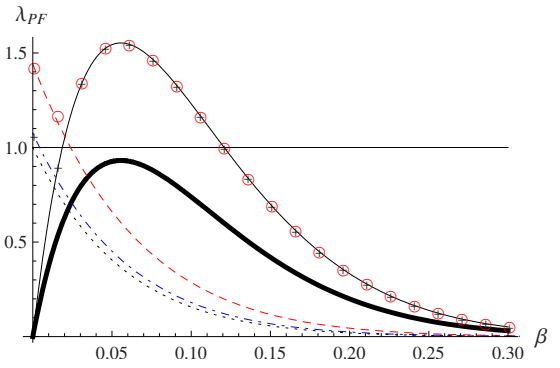

FIG. 1. (Color online) We show here the dependence of the Perron-Frobenius eigenvalue λ_{PF} of $S^{-1}E_1^{\tau}$ on the global mutation rate β for selective advantage $k=0.5$, period $\tau=18$, and genome length $\ell=100$. The data points located as empty dots (colored red in the online version) are numerically calculated results for matrix size $M=20$, whereas the + signs (black) are numerical results for $M=100$. The light solid line is the plot of approximation (23) and the thicker solid line is the plot of the right-hand side of Eq. (9). The dashed line (red), the dotted-dashed line (blue), and the dotted line (black) represent, respectively, the first-order perturbative results for $M=20$, $M=100$, and $M=\infty$.

We have included in Fig. 1 a plot of the third eigenvalue of $S_{\mathrm{NS}}E_{\mathrm{NS}}^{\tau}$ given by Eq. (9). Notice that it does not approximate the numerically calculated λ_{PF} and is smaller than 1 for all β's. The set of parameter values in Fig. 1 exemplifies one case in which our analysis shows the existence of a quasispecies for some range in β, whereas the analysis in [15,21] does not.

In Fig. 2 we show location in the complex plane of all eigenvalues of matrix $S^{-1}E_1^{\tau}$ for parameters $\beta=0.1$, $k=1.1$, $\tau=10$, $\ell=100$, and $M=20$ or $M=100$. Notice in Fig. 2 that λ_{PF} is noticeably larger than 1 and practically independent of M, whereas all other eigenvalues, although noticeably dependent on M, have much smaller absolute values. Due to this fact, any non-negative initial condition $p(0)$ will evolve so as to become after a few cycles of τ generations practically parallel to v_{PF}. Convergence to the steady-state frequencies is thus rather fast.

In Fig. 3 we show logarithmic plots of the components of v_{PF}. We take $k=1.2$, $\tau=10$, $\ell=100$, and $M=60$ and two val-

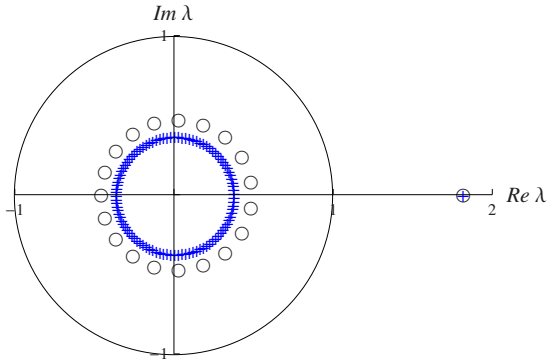

FIG. 2. (Color online) Location in the complex plane of the eigenvalues of $S^{-1}E_1^{\tau}$ for parameters $\beta=0.1$, $k=1.1$, $\tau=10$, and $\ell=100$. Empty dots (black) are for $M=20$ and + signs (blue) are for $M=100$. The unit circle is also plotted for convenience.

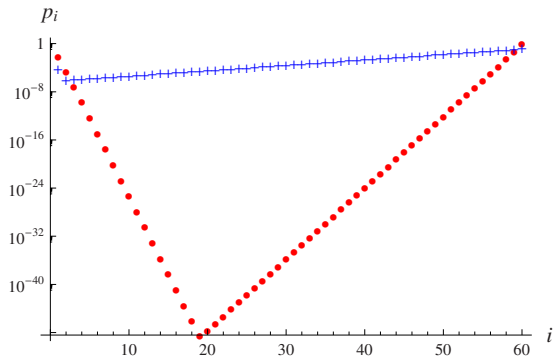

FIG. 3. (Color online) Logarithmic plots of the components of the Perron-Frobenius eigenvector of $S^{-1}E_1^\tau$. The points located by + signs (blue) refer to $\beta=0.017$ and points located by dots (red) refer to $\beta=0.28$. The other parameters values are $k=1.2$, $\tau=10$, $\ell=100$, and $M=60$ in both cases. In both cases the eigenvectors are normalized such that the sum of their components equals 1.

ues for β: 0.017, just above the lower threshold for the quasispecies existence, and 0.28, just below the upper threshold. In both cases the eigenvectors are normalized such that the sum of their components equals 1. We see that in both cases the largest component is the last one. We also see that components seem to decay approximately exponentially at the beginning, attain a minimum, and then increase exponentially. Interestingly, the rate of decay of the first components is larger than the rate of growth of the last components, with this difference being much larger for the smaller values of β.

Figure 3 shows thus two "snapshots" of quasispecies at a time in which σ_1 is the master sequence. So, the vector components after the first refer to sequences which will become master in a few cycles, and the last component and its neighbors refer to sequences which have been masters not too many cycles ago. This explains the "asymmetry" in the snapshots as the fact that populations in the former master sequences take some time to decay. In the analysis of [15,21], this "memory effect" is not taken into account because of the separation of all individuals into error classes.

Figure 3 provides also information on the genetic variability of the population. Such interesting information is not supplied by the coarser analysis of [15,21].

Although details are to be left to [31], it is important here to stress that it is possible to derive good approximate formulas for λ_{PF} and explain where these formulas come from. For any $n\times n$ matrix A with elements $a_{i,j}$, the *Collatz-Wielandt function* f_A is defined as

$$f_A(x)=\min_{x_i\neq 0}\frac{(Ax)_i}{x_i}, \tag{20}$$

where x are vectors in $\mathbb{R}^n$, $(Ax)_i$ and x_i are the ith components of the vectors Ax and x, and the minimum is taken over all the indices i such that $x_i\neq 0$. A common method for proving the PF theorem for irreducible matrices [30] goes through showing that f_A attains a maximum over the set $\mathbb{P}^n$ of non-negative vectors in $\mathbb{R}^n$, that this maximum is exactly the PF eigenvalue, and the vector x which maximizes f_A is the corresponding eigenvector. Then for any $x\in\mathbb{P}^n$, the PF eigenvalue λ_{PF} of A is such that $\lambda_{\mathrm{PF}}\geq f_A(x)$.

Let e_k be the unit vector with all components equal to zero with the exception of the kth. We may obtain lower bounds for λ_{PF} by taking, for any pair i,j of indices with $i<j$, vectors of the form $x_{i,j}(\alpha)=\alpha e_i+(1-\alpha)e_j$ and calculating α, so that $f_A(x_{i,j}(\alpha))$ is maximized. It results that for any irreducible non-negative matrix A and any pair of indices $i<j$,

$$r_{i,j}=a_{j,j}+\frac{\alpha_{i,j}}{1-\alpha_{i,j}}a_{j,i} \tag{21}$$

is a lower bound to the PF eigenvalue, where

$$\alpha_{i,j}=\frac{a_{i,i}-a_{j,j}-2a_{i,j}+\sqrt{(a_{i,i}-a_{j,j})^2+4a_{i,j}a_{j,i}}}{2[(a_{i,i}-a_{j,j})+(a_{j,i}-a_{i,j})]}. \tag{22}$$

In the case $A=S^{-1}E_1^\tau$ this upper bound with $i=1$ and $j=M$ is also an excellent approximation for λ_{PF} for not too small values of β. In [31] we will explain why this is so. Also in [31] we write elements of E_1^τ as a sum of contributions from directed graphs and obtain good approximations to them by selecting suitable terms in these expansions. Using these results in Eqs. (21) and (22) with $i=1$, $j=M$ we finally get

$$\lambda_{\mathrm{PF}}\approx\frac{(1+k)^\tau(2+k)}{k\ell}\beta(1-\beta)^{\tau-1}. \tag{23}$$

In Fig. 1 we plot the right-hand side of the above equation and confirm that it is indeed a good approximation for λ_{PF} unless β is very close to zero.

It is remarkable that the above approximation has an identical dependence on β as Eq. (9), but the multiplicative constant is different. As both expressions coincide asymptotically as $k\to\infty$, we conjecture that the result of [15,21] should be exact in this asymptotic limit.

On the other hand, not only Eq. (23) fails for small β, but also there is no result in [15,21] for this region. The reader may notice that E_1 becomes diagonal if we take limits $\mu\to 0$, $\ell\to\infty$ such that β remains constant, and it follows that $S^{-1}E_1^\tau$ is exactly diagonalizable in this limit. It is then possible to approximate λ_{PF} to first-order perturbation theory in parameter $\epsilon=\ell^{-1}$. We refer to [31] for the details. The result is

$$\lambda_{\mathrm{PF}}=(1-\beta)^\tau(1+k)^{\tau/M}+\frac{1}{\ell}\lambda_1+O(1/\ell^2), \tag{24}$$

where

$$\lambda_1=\beta(1-\beta)^{\tau-1}\left\{\tau\left(1-\frac{2}{M}\right)[1+(1+k)^{2\tau/M}]+\frac{k+2}{Mk}[(1+k)^\tau-(1+k)^{-\tau}]\right\}. \tag{25}$$

The right-hand side of Eq. (24) with the $O(1/\ell^2)$ term dropped is also plotted in Fig. 1 for $M=20$ and $M=100$. We see that these plots agree well with the corresponding numerical results for small β.

The reader should also notice that the limit $M \to \infty$ can be taken in Eq. (24), and we end up with the perturbative approximation

$$\lambda_{\mathrm{PF}} = (1-\beta)^{\tau}\left(1 + \frac{2\tau\beta}{\ell(1-\beta)}\right) + O\left(\frac{1}{\ell^2}\right), \tag{26}$$

which is smaller than 1 for β close to zero. This justifies the claim we had made before that the small intervals close to $\beta=0$ in which $\lambda_{\mathrm{PF}}>1$ are in fact an artifact of taking finite (and small) values of M.

V. DISCUSSION AND CONCLUSIONS

We have analyzed the same moving SPL as in the seminal papers by Nilsson and Snoad [15] and Ronnewinkel *et al.* [21]. Whereas the results in the above papers can be seen, as we did in Sec. II, as resulting from exact diagonalization of a 3×3 matrix, the results in the present paper depend on the possibility of approximately diagonalizing matrix $S^{-1}E_1^{\tau}$ of order M. In principle, M should be equal to the number of cycles of τ generations it takes until a sequence which had been master becomes master again. Typically, this number is of order 2^{ℓ} cycles, with ℓ being the size of the genome.

The existence or nonexistence of a region in parameter β in which a quasispecies exists, as well as the thresholds that bound this region, was related in Sec. III to the PF eigenvalue λ_{PF} of matrix $S^{-1}E_1^{\tau}$. In Sec. IV we obtained two formulas which approximate very well λ_{PF}.

Formula (23), which does that for larger values of β is already independent of M. This independence stems from the fact that the elements of matrix $S^{-1}E_1^{\tau}$, used in deriving Eq. (23), depend very slightly on M. This is by its turn a consequence of the fact that if we write in components $p(t+1)=E_1 p(t)$, we will find a discretized version of the heat equation with losses

$$\frac{\partial u}{\partial t} - D\frac{\partial^2 u}{\partial x^2} + \gamma u(x,t) = 0,$$

in which $\gamma>0$ is so large that the heat loss term dominates the diffusion term. In such a case, what happens at a point x depends only on its immediate neighborhood and is almost independent of the rest of the system, so that the length M of the heat conducting region does not matter. Figure 1 shows indeed that the numerically calculated λ_{PF}'s are well approximated by Eq. (23) and almost independent of M.

Figure 1 shows also that for small β the numerical values of λ_{PF} are well approximated by Eq. (24). In this region, however, unrealistically small values of M such as 20 or 100 do not provide a good approximation to the biologically relevant limit $M\to\infty$. Anyway, taking this limit in Eq. (24) is easy, and we obtain Eq. (26) which should be a good approximation for λ_{PF} for values of M of order 2^{ℓ} and β very close to zero.

The results obtained so far should be valid only if $\sigma_1,\ldots,\sigma_M$ are all distinct sequences. But, as we have already commented, this is not a typical situation if M is as large as 2^{ℓ}. We should in fact expect two-step backjumps of the master sequence occurring in average each ℓ cycle of τ generations. As shown in Fig. 13 of [21] such backjumps have the effect of increasing the population at the master sequence beyond the value it would have if there were no backjumps, but this effect dies out quickly. This phenomenon can be explained by noticing that in typical situations, such as those depicted in Fig. 2, λ_{PF} is much larger than all other eigenvalues of $S^{-1}E_1^{\tau}$ in absolute value. Thus, any perturbation in the dynamics, such as occurrence of a backjump, should be exponentially damped in a few cycles of τ generations, i.e., in a time much smaller than the typical time between two backjumps. As a consequence, backjumps cannot alter the long-time behavior of the system in a typical situation in which λ_{PF} is largely dominant. An instance in which backjumps may be important is for small β, because in this case λ_{PF} may be close in absolute value to the other eigenvalues, but we did not analyze further this situation.

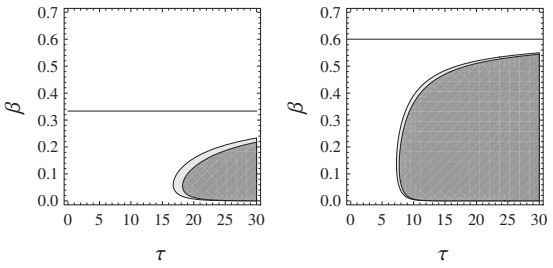

FIG. 4. Phase diagrams showing for two values of k (left diagram is for $k=0.5$ and right diagram is for $k=1.5$) the regions of quasispecies existence in the analysis of the present paper and in the analyses of [15,21]. The dark gray region is the quasispecies region as in [15,21], while our own analysis indicates the same region augmented by the light gray strip. The horizontal lines included for reference represent the error threshold in the static SPL case.

The relationship between our results and those of [15,21] is illustrated also in the phase diagrams of Fig. 4, in which we show for two values of k the regions in the (τ,β) plane in which a quasispecies exists in both analyses. As it can be seen, our region for the existence of a quasispecies properly contains that of [15,21]. But the difference is smaller for larger values of k, and both results coincide asymptotically when $k\to\infty$, as an inspection of Eqs. (9) and (23) shows.

We believe that the enlargement of the quasispecies region with respect to [15,21], mainly for large values of β, as seen in Fig. 4, has two main sources. One is the fact illustrated in Fig. 3 that sequences which have already been masters are more populated than sequences at the same HD which have never been masters. We believe that the high population of the sequence in Γ_1 which has just ceased to be master helps increasing the population in the sequence in Γ_1 which will be the next master sequence, mainly if β is large. The second source is the fact that we need not suppose Eq. (6). A third source might be the fact that we do consider some backmutations in our analysis, but we do not believe this is so important. In fact we considered such backmutations only because equations were more symmetric, and it did not introduce any technical difficulties.

Our analysis not only confirms the validity of the error thresholds of [15,21] asymptotically as $k\to\infty$, but also provides more correct results for small k. This numerical differ-

ence could be relevant when considering the use of mutagen drugs as antiviral therapy, as suggested, e.g., in [28]. Moreover, we add more detailed information on the genetic variability of the quasispecies, as can be seen in Fig. 3, not available in the results of [15,21].

A different result from our analysis is that λ_{PF} is always larger than 1 for β in an interval $[0,\beta_{ll})$, as can be seen from either Fig. 1 or Eq. (24). We showed that β_{ll} tends to zero when $M\to\infty$, so that this tiny interval of the existence for a quasispecies vanishes in the relevant limit for the situation examined here. But in a situation in which the environment fluctuates periodically and only a small number of genotypes $\sigma_1,\ldots,\sigma_M,\sigma_{M+1}$, with $\sigma_{M+1}\equiv\sigma_1$, and $d(\sigma_i,\sigma_{i+1})=1$, with $i=1,2,\ldots,M$, ever become master sequences, the interval $[0,\beta_{ll})$ represents an additional interval where a quasispecies exists, even for values of k,τ,ℓ for which the other existence interval (β_l,β_u) does not exist. We have not explored this possibility further.

We indicate here two other possible uses of the tools developed here in future works. One would be to extend, as in [25], the analysis to a periodically moving SPL in which the master sequence after a shift is at HD d of the previous one. If $d\ll M$, this can be easily done by replacing matrix $\hat{S}$ in Eq. (16) with its counterpart for a shift d units apart. From the point of view of numerical calculations, there seems to be no problem, but we do not know yet how to obtain results analogous to Eq. (23). A second possible extension of the tools developed here could be the study of a dynamic SPL in which τ is random. For example, it can be easily seen that if τ does not fluctuate too much, so that the point (β,τ) always falls in the dark region in Fig. 4, a quasispecies will exist. Of course the problem becomes more difficult if this condition does not hold.

ACKNOWLEDGMENTS

The author acknowledges CNPq, an agency of the Brazilian Ministry for Science and Technology, for support and the Mathematics Department of the Università degli Studi di Milano for hospitality during preparation of this work. Thanks are due to Giuseppe Gaeta for many discussions and suggestions and the anonymous referees for clarifying questions and suggestions which helped improving this work.

[1] M. Eigen, Naturwiss. **58**, 465 (1971).
[2] E. Tannenbaum and E. I. Shakhnovich, Phys. Life. Rev. **2**, 290 (2005).
[3] J. F. Crow and M. Kimura, *An Introduction to Population Genetics Theory* (Harper & Row, New York, 1970).
[4] J.-M. Park and M. W. Deem, J. Stat. Phys. **125**, 971 (2006).
[5] E. Muñoz, J.-M. Park, and M. W. Deem, J. Stat. Phys. **135**, 429 (2009).
[6] L. Peliti, e-print arXiv:cond-mat/9712027.
[7] E. Baake and W. Gabriel, Annu. Rev. Comput. Phys. **VII**, 203 (2000).
[8] B. Drossel, Adv. Phys. **50**, 209 (2001).
[9] I. Leuthusser, J. Stat. Phys. **48**, 343 (1987).
[10] P. Tarazona, Phys. Rev. A **45**, 6038 (1992).
[11] E. Baake, M. Baake, and H. Wagner, Phys. Rev. Lett. **78**, 559 (1997).
[12] D. Saakian and C.-K. Hu, Phys. Rev. E **69**, 021913 (2004).
[13] D. B. Saakian and C.-K. Hu, Phys. Rev. E **69**, 046121 (2004).
[14] D. Saakian and C. Hu, Proc. Natl. Acad. Sci. U.S.A. **103**, 4935 (2006).
[15] M. Nilsson and N. Snoad, Phys. Rev. Lett. **84**, 191 (2000).
[16] M. Nilsson and N. Snoad, Bull. Math. Biol. **64**, 1033 (2002).
[17] M. Nilsson and N. Snoad, Phys. Rev. E **65**, 031901 (2002).
[18] C. O. Wilke, C. Ronnewinkel, and T. Martinetz, Phys. Rep. **349**, 395 (2001).
[19] C. O. Wilke and C. Ronnewinkel, Physica A **290**, 475 (2001).
[20] C. O. Wilke, C. Ronnewinkel, and T. Martinetz, in *Proceedings of the ECAL'99*, edited by H. Lund and R. Kortmann (Springer-Verlag, Heidelberg, 1999).
[21] C. Ronnewinkel, C. O. Wilke, and T. Martinetz, in *Theoretical Aspects of Evolutionary Computing*, edited by L. Kallel, B. Naudts, and A. Rogers (Springer-Verlag, Heidelberg, 2001).
[22] Y. Brumer and E. I. Shakhnovich, Phys. Rev. E **71**, 031903 (2005).
[23] C. Kamp and S. Bornholdt, Phys. Rev. Lett. **88**, 068104 (2002).
[24] P. Gorodetsky and E. Tannenbaum, Phys. Rev. E **77**, 042901 (2008).
[25] M. Ancliff and J.-M. Park, Phys. Rev. E **80**, 061910 (2009).
[26] J.-M. Park and M. Ancliff, J. Korean Phys. Soc. **56**, 891 (2010).
[27] M. Ancliff and J.-M. Park, J. Korean Phys. Soc. **54**, 2440 (2009).
[28] M. Eigen, Proc. Natl. Acad. Sci. U.S.A. **99**, 13374 (2002).
[29] J.-M. Park, E. Muñoz, and M. W. Deem, Phys. Rev. E **81**, 011902 (2010).
[30] H. Y. Minc, *Nonnegative Matrices* (John Wiley and Sons, New York, 1988).
[31] A. G. M. Neves (unpublished).