Evolutionary dynamics of the most populated genotype on rugged fitness landscapes

Kavita Jain

Department of Physics of Complex Systems, Weizmann Institute of Science, Rehovot 76100, Israel

We consider an asexual population evolving on rugged fitness landscapes which are defined on the multi-dimensional genotypic space and have many local optima. We track the most populated genotype as it changes when the population jumps from a fitness peak to a better one during the process of adaptation. This is done using the dynamics of the shell model which is a simplified version of the quasispecies model for infinite populations and standard Wright-Fisher dynamics for large finite populations. We show that the population fraction of a genotype obtained within the quasispecies model and the shell model match for fit genotypes and at short times, but the dynamics of the two models are identical for questions related to the most populated genotype. We calculate exactly several properties of the jumps in infinite populations, some of which were obtained numerically in previous works. We also present our preliminary simulation results for finite populations. In particular, we measure the jump distribution in time and find that it decays as t^{-2} as in the quasispecies problem.

I. INTRODUCTION

The question of mode in evolution, especially in the context of speciation, has engaged the attention of many evolutionists for over a century and continues to do so [1]. The issue is whether evolution occurs by smooth gradual changes (gradualism) as put forwarded by Darwin, or sudden large jumps (punctuationism) [2,3]. Some examinations of fossil record indicate that new species can arise by either mode, or even by a combination of the two [4]. However, a complete and unambiguous answer is hard to obtain at the level of macroevolution due to the incompleteness and irreproducibility of the fossil data. In the last decade or so, researchers have become interested in carrying out long-term evolution experiments in the laboratory [5]. Typically, one starts with a microbial population maladapted to a given environment such as a colony of starving bacteria, and track its evolutionary trajectories for thousands of generations as it undergoes adaptive changes. The results of such experiments have been found to be consistent with both modes of evolution. For instance, large populations of RNA virus starting from a low fitness ancestor have been seen to gain fitness in a continuous manner [6]. On the other hand, the fitness of bacteria *E. Coli* [7,8] and RNA virus ϕ_6 [9] show a punctuated pattern of evolution.

Theoretically, these results are understood using the concept of fitness landscape [10–12] defined as a map from the genotypic space into the real numbers. If the fitness landscape is smooth and single peaked, starting from a low fitness state the population fitness increases gradually until it has reached the peak value [13]. The dynamics are different for a population moving on a rugged fitness landscape with multiple peaks [14]. In this case, the population fitness increases smoothly until the population encounters a local fitness peak where it gets trapped as a better peak is separated by a fitness valley. The population thus enters the stasis phase and waits until a favorable mutation allows it to shift to a better peak where it can again get localized and so on. Thus, the dynamics alternate between periods of *stasis* and

rapid changes in fitness when the population *jumps* from a peak to an even better one. An example of this behavior is shown in Fig. 1.

That the fitness landscape underlying the evolutionary process is multipeaked is supported by several experiments [7,9,15–18]. Besides, these landscapes include biologically important and ubiquitous epistatic interactions [19] as the contribution of individual genes to the fitness of the genotype is not independent [14]. The class of landscapes considered in this work are maximally rugged and characterized by strong selection [20]. In statistical physics, such rugged landscapes with a large number of local optima have appeared in the context of spin glass theory. Examples include the random energy model [21] and the Sherrington-Kirkpatrick model [22] where the energy of a spin configuration plays the role of genotypic fitness and the metastability in the spin glass dynamics can be viewed as punctuated equilibrium [23].

We are interested in the statistics of jumps which occur when the population fitness changes rapidly. As illustrated in Fig. 1, an unambiguous way of defining a jump at time t is the change in the fitness or identity of the *most populated genotype* at t. The problem of such leadership changes arises in several other contexts such as change in the position of the optimal end point of the directed polymer [24], highest degree node in a growing network [25,26], and velocity and position of the front particle in a single-file system [27,28]. In the following sections, before analyzing the realistic case of finite populations, we will study the infinite population limit in detail. We consider the dynamics of Eigen's model [29] that describes the population dynamics of self-replicating molecules which at low mutation rates and large times form a dynamic and heterogeneous quasispecies consisting of the fittest genotype and its closely related mutants. The existence of such an error threshold is a generic result seen in both simple and complex fitness landscapes, and has been reviewed in Ref. [30] (for recent related works, see Refs. [31–34]). The focus of this paper is, however, the *dynamics* of the quasispecies model. We show that the population fraction at a genotype obtained within a simplified shell
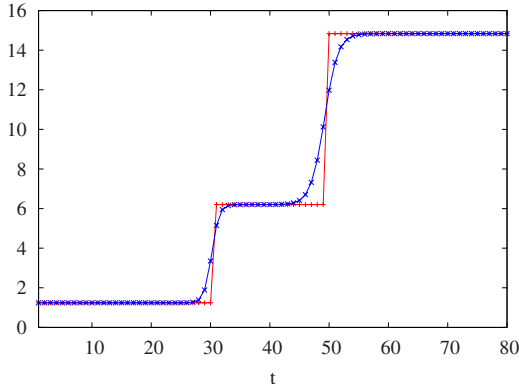
FIG. 1. (Color online) Punctuated change in the population fitness ($\times$) and the fitness of the most populated genotype ($+$) in a single realization of a maximally rugged fitness landscape. Here $L=6$, $N=2^{14}$, $\mu=10^{-6}$, $p(W)=W^{-2}$.

model [35] (also see Ref. [36]) of quasispecies dynamics is good only for highly fit sequences and at short times. However, the behavior of the most populated genotype is captured correctly by the shell model. We calculate exactly some properties of the jumps (as defined above) within the shell model which were obtained numerically in previous works [35,37,38]. Specifically, we show that the jump distribution decays as $1/t^2$ in time and $1/\sqrt{k}$ as a function of the distance k from the starting genotype for a class of fitness landscapes.

The basic difference between a finite and infinite population is that while the former has a finite mutational spread in the genotypic space, all the mutants are available at all times in the quasispecies limit. Due to this reason, at large times, a finite population gets trapped at a local fitness peak and the jump probability is governed by the rate of stochastic tunneling [39,40] which allows the population to cross the fitness valley via few low fitness mutants. In the quasispecies model, on the other hand, the reproduction or selection plays a more important role than the production of mutants and a jump occurs when the population at a fit genotype overtakes the less fit one. Although the underlying physical processes responsible for a jump are different for finite and infinite populations, we find that the jump distribution is robust in that it decays as $1/t^2$ in both the cases.

The plan of the paper is as follows. In the next section, we define a class of mutation-selection models. Section III discusses the dynamics of the quasispecies model and the shell model. We calculate exactly the statistics of jumps within the shell model in Sec. IV. These results for infinite population limit are compared with those obtained in simulations of finite population in Sec. V. Finally, we conclude with a summary of our results.

II. MODELS

We consider a haploid, asexual population of size N each of whose constituents carry a string $\sigma=\{\sigma_1,\ldots,\sigma_L\}$ of length L where σ_i can assume $\ell \geq 2$ values. The string σ can represent a genetic sequence ($\ell=4$), protein ($\ell=20$), or a sequence of L loci with ℓ alleles [30]. For simplicity, we will

deal with binary sequences for which $\ell=2$ and $\sigma_i=0$ or 1. The environment of the population is represented by a fitness landscape which is obtained by associating a non-negative real number $W(\sigma)$ to each sequence. In this article, we consider fitness $W(\sigma)$ to be a random variable chosen independently from a common distribution $p(W)$. This generates a maximally rugged fitness landscape with an exponentially large number (in L) of local maxima [14,41] and strong interactions amongst the loci [20]. The population evolves on this fitness landscape in discrete time via selection and mutation, and we ignore other factors responsible for genetic mixing such as recombination. Our model is thus applicable to microorganisms like *E. Coli* and Hepatitis C virus which have zero or very low recombination rates [42].

Consider a population well adapted to a given environment localized around a peak of the fitness landscape. A change in the environment brings about a change in the fitness landscape, and the population will typically find itself in a fitness valley. We consider the adaptation process of the population starting from such an initial condition. The population fraction $X(\sigma,t)$ of sequence σ at time t is iterated using Wright-Fisher dynamics defined as follows. In each generation, an offspring selects a parent p with probability $W_p(\sigma)/N\langle W\rangle$, where $W_p(\sigma)$ is the fitness of the parent p with sequence σ and $\langle W\rangle=\Sigma_{\sigma'}W(\sigma')X(\sigma',t)$ is the average fitness of the population. Then the probability $P(n)$ that parent p has n offspring in one generation is given as

$$P(n) = \binom{N}{n}\left(\frac{W_p(\sigma)}{N\langle W\rangle}\right)^n\left(1-\frac{W_p(\sigma)}{N\langle W\rangle}\right)^{N-n}.$$

This implies that the average number of offspring produced equals $W(\sigma)/\langle W\rangle$ and hence the fitness $W(\sigma)$ has the physical meaning that it is proportional to the number of descendants produced per generation. Further, as the relative variance in the offspring number decays as $1/N$, it follows that the offspring number fluctuates from one generation to another for finite population, but one can ignore these fluctuations arising due to random sampling for $N\to\infty$. After the reproduction process, point mutations are introduced independently at each locus of the sequence σ' with probability μ per generation. Thus, a sequence σ is obtained via mutations with the probability

$$p_{\sigma\leftarrow\sigma'} = \mu^{d(\sigma,\sigma')}(1-\mu)^{L-d(\sigma,\sigma')}, \tag{1}$$

where the Hamming distance $d(\sigma,\sigma')$ is the number of point mutations in which the sequences σ and σ' differ.

Since the selection process is not stochastic for infinite populations, the population frequency also does not fluctuate and one can work with the average frequency $\mathcal{X}(\sigma,t)=\langle X(\sigma,t)\rangle$ where the averaging is over all realizations of the sampling process. This leads to a deterministic nonlinear equation for the fraction $\mathcal{X}(\sigma,t)$ [20],

$$\mathcal{X}(\sigma,t+1) = \frac{\sum_{\sigma'} p_{\sigma\leftarrow\sigma'}W(\sigma')\mathcal{X}(\sigma',t)}{\sum_{\sigma'} W(\sigma')\mathcal{X}(\sigma',t)}, \tag{2}$$

where the denominator is the normalization constant.

III. DYNAMICS OF THE QUASISPECIES AND SHELL MODELS

The focus of this section is the quasispecies model and the related shell model. In the following, we will mainly work with the unnormalized population $\mathcal{Z}(\sigma,t)$ defined as [30]

$$\mathcal{Z}(\sigma,t) = \mathcal{X}(\sigma,t)\prod_{\tau=0}^{t-1}\sum_{\sigma'} W(\sigma')\mathcal{X}(\sigma',t) \quad (3)$$

in terms of which the nonlinear evolution (2) reduces to the following linear iteration:

$$\mathcal{Z}(\sigma,t+1) = \sum_{\sigma'} p_{\sigma\leftarrow\sigma'} W(\sigma')\mathcal{Z}(\sigma',t). \quad (4)$$

Since at the start of the adaptation process the population finds itself at a low fitness genotype, we start with the initial condition $\mathcal{X}(\sigma,0) = \mathcal{Z}(\sigma,0) = \delta_{\sigma,\sigma^{(0)}}$, where $\sigma^{(0)}$ is a randomly chosen sequence. For small mutation probability μ ($\sim 10^{-3} - 10^{-10}$) as seen in various asexual microbes [30,43], after one iteration we have

$$\mathcal{Z}(\sigma,1) = \mu^{d(\sigma,\sigma^{(0)})} W(\sigma^{(0)}). \quad (5)$$

Thus each sequence gets populated in one generation with a fraction which is same for all the sequences in a *shell* of constant Hamming distance $d(\sigma,\sigma^{(0)})$ from the initial sequence $\sigma^{(0)}$ [35]. Numerical simulations of Ref. [35] showed that dynamical properties involving the most populated genotype, such as the distribution of evolution times and number of jumps, are very well described by a simplified model which ignores mutations for further evolution and allows the population at each sequence to grow with its own fitness. Thus, within the shell model, the population $\mathcal{Z}(\sigma,t) \sim \mu^{d(\sigma,\sigma^{(0)})} W^t(\sigma)$ for $t>1$.

Here we provide an analytical understanding of the quasispecies model leading to the shell model. We will find that the expression for $\mathcal{Z}(\sigma,t)$ in the shell model is a good approximation for sequences with high fitness and at short times. For $t>1$, consider (4) for a sequence σ in the shell d centered about $\sigma^{(0)}$ and at a Hamming distance $d(\sigma,\sigma^{(0)})$ from the center. The sum on the right-hand side of (4) has three kind of terms: (i) sequence σ does not mutate (i.e., $\sigma' = \sigma$ in the sum) so that its contribution $\propto \mu^{d(\sigma,\sigma^{(0)})}$, (ii) a mutation occurs in $\sigma' \in I$ where I is the set of σ''s lying inside shell d that satisfy $d(\sigma,\sigma') + d(\sigma',\sigma^{(0)}) = d(\sigma,\sigma^{(0)})$ resulting in $\mu^{d(\sigma,\sigma^{(0)})}$ dependence, and (iii) a mutation occurs either in the sequences in the inner shells that do not belong to set I or, in the sequences in and outside shell d giving a term of order $\mu^{d(\sigma,\sigma^{(0)})+1}$ and higher. In order to obtain $\mathcal{Z}(\sigma,t)$ to order $\mu^{d(\sigma,\sigma^{(0)})}$, we can neglect the last contribution and iterate the population according to

$$\mathcal{Z}(\sigma,t+1) = W(\sigma)\mathcal{Z}(\sigma,t) + \sum_{\sigma' \in I} \mu^{d(\sigma,\sigma')} W(\sigma')\mathcal{Z}(\sigma',t). \quad (6)$$

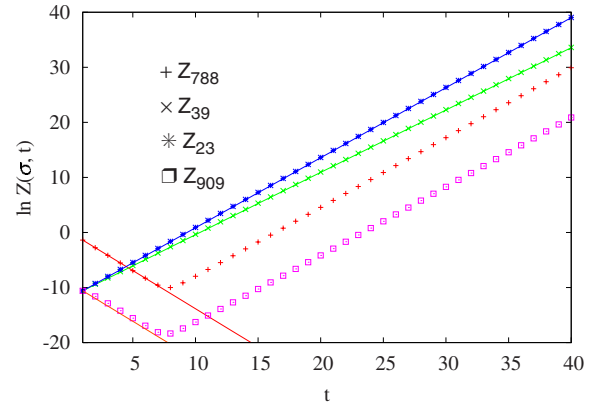The above equation is still coupled, but one can make further simplifications by proceeding as follows. Let us first

FIG. 2. (Color online) Comparison of the logarithmic population fraction $\ln \mathcal{Z}(\sigma,t)$ obtained by iterating exact equation (4) (shown by points) and the shell model dynamics (7) and (9) (shown by lines) for $\mu = 10^{-4}$. The fitness W is chosen from a common exponential distribution for a sequence of length $L=10$. The subscript in $\mathcal{Z}$ gives the fitness rank where the largest fitness carries rank 0 and the population initially starts at a sequence with rank 788. All the other sequences are at Hamming distance 1 from $\sigma^{(0)}$.

consider the zeroth shell, i.e. sequence $\sigma^{(0)}$ for which we immediately have

$$\mathcal{Z}(\sigma^{(0)},t) = W^t(\sigma^{(0)}). \quad (7)$$

For the sequences in the first shell for which $d(\sigma,\sigma^{(0)}) = 1$, the sequence $\sigma^{(0)}$ is the only member of set I and we have

$$\mathcal{Z}(\sigma,t+1) = W(\sigma)\mathcal{Z}(\sigma,t) + \mu W^{t+1}(\sigma^{(0)}), \quad d(\sigma,\sigma^{(0)}) = 1. \quad (8)$$

Defining $\mathcal{Z}(\sigma,t) = W^t(\sigma)z(t)$, we obtain a simple difference equation for $z(t)$ which can be iterated to give

$$z(t) = \mu r + \mu r^2 \frac{1-r^{t-1}}{1-r}, \quad r = \frac{W(\sigma^{(0)})}{W(\sigma)}.$$

If $W(\sigma) > W(\sigma^{(0)})$, the population $\mathcal{Z}(\sigma,t)$ grows exponentially fast with a rate equal to its own (log) fitness whereas in the opposite case, this growth rate is that of the initial sequence. For the sequences in the first shell, one therefore obtains

$$\mathcal{Z}(\sigma,t) = \begin{cases} \mu W(\sigma^{(0)})W^{t-1}(\sigma), & r < 1 \\ \mu W^t(\sigma^{(0)}), & r > 1. \end{cases} \quad (9)$$

For the succeeding shells, one can work out the population $\mathcal{Z}(\sigma,t)$ in a manner similar to above and find that $\mathcal{Z}(\sigma,t) \sim \mu^{d(\sigma,\sigma^{(0)})} W^t(\sigma)$ if the fitness $W(\sigma)$ is larger than that of all the sequences in set I. Such a random variable whose value is larger than all the ones preceding it defines a *record* [37,38]. If the fitness $W(\sigma)$ is not a record, then the growth rate is given by the largest (log) fitness of sequences in set I, i.e. the last record value.

A comparison of the results obtained using the above approximations with the exact iteration of (4) is shown in Fig. 2 for the initial sequence and some representative sequences in the first shell. The disagreement at large times occurs be-

cause (7) and (9) are obtained to leading order in μ, and after some time as the population at the outer shells grow, the next order terms in μ start contributing. For the zeroth shell, this happens at time τ when the next order terms $\mu \mathcal{Z}(\sigma, \tau)$ due to sequences in shell 1 become as large as the lowest order term given by (7), i.e. $W^\tau(\sigma^{(0)}) \sim \mu^2 W^\tau(\sigma^*)$ where $W(\sigma^*) > W(\sigma^{(0)})$ is the largest fitness in shell 1. Plugging in the fitness values of relevant sequences in this expression for the fitness landscape in Fig. 2, we obtain $\tau \approx 8$ in good agreement with the exact iteration. After time τ, the population at $\sigma^{(0)}$ grows with the fitness $W(\sigma^*)$ until shell 2 starts contributing. This argument can be generalized to higher shells straightforwardly. For instance, the correction to (9) is of order μ^3 which arises either due to the sequences in the first shell but two mutations away from σ or the sequences in the second shell that are one mutational distance away. The larger of the two contributions can be then used to estimate the time at which the growth rate changes. This process of slope changes of logarithmic population goes on until the globally fittest sequence $\sigma^{(f)}$ becomes most populated after which the population $\mathcal{Z}(\sigma, t) \sim \mu^{d(\sigma, \sigma^{(f)})} W^t(\sigma^{(f)})$.

In Fig. 2, the initial sequence with rank 788 remains most populated until ~ 5 time steps after which the sequence ranked 23 overtakes it and becomes the next leader. We are interested in such leadership changes and will use the shell model for this purpose as it correctly captures the dynamical behavior of the most populated genotype. Recall that in the shell model, the population at each sequence grows with its own intrinsic fitness for all $t > 1$. These population dynamics are different from quasispecies in which the population of a sequence whose fitness is not a record grows with the fitness of the last record (in inner shells) at short times and the growth rate of the sequences change as the leading genotype changes in course of time. However, since the population at such sequences is always at least order μ smaller than that at the leading genotype, the dynamics of the most populated sequence are not affected if the population at such sequences is also allowed to grow with their respective fitness.

After these considerations, we arrive at the shell model in which the logarithmic population obeys the following linear time evolution:

$$\ln \mathcal{Z}(\sigma, t) = \ln W(\sigma^{(0)}) - |\ln \mu| d(\sigma, \sigma^{(0)}) + \ln W(\sigma)(t-1). \tag{10}$$

Calling $F(\sigma) = \ln W(\sigma)$ and rescaling time by $|\ln \mu|$, we have

$$E(\sigma, t) = -d(\sigma, \sigma^{(0)}) + F(\sigma)t, \tag{11}$$

where we have absorbed the extraneous factors in the definition of logarithmic population $E(\sigma, t)$. Since the population at $t = 1$ is same for sequences at constant $d(\sigma, \sigma^{(0)}) = k$, only the sequence with the largest fitness in shell k needs to be considered [35]. If the logarithmic fitness F is distributed according to the distribution $p(F)$, then the largest fitness $F(k)$ in shell k which is the maximum of $\alpha_k = \binom{L}{k}$ variables is distributed independently but nonidentically with distribution
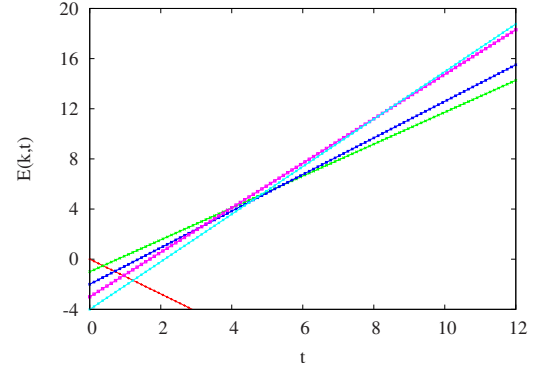

FIG. 3. (Color online) Shell model dynamics for population $E(k, t) = -k + F(k)t$ for $k = 0, \dots, 4$ for the landscape in Fig. 2. The sequence with the largest fitness in shell 2 gets bypassed by the corresponding sequence in shell 3.

$$p_k(F) = \alpha_k p(F) \left[\int_{F_{\min}}^{F} dF' p(F') \right]^{\alpha_k - 1}, \tag{12}$$

where $F_{\min}$ is the lower support of the distribution $p(F)$.

IV. STATISTICS OF JUMPS IN THE SHELL MODEL

In the following, we will consider the shell model dynamics defined as

$$E(k, t) = -k + F(k)t, \tag{13}$$

where $F(k)$ and $E(k, t)$, respectively, are the fitness and population of the sequence with the largest fitness in shell $k = 0, \dots, L$. Figure 3 shows the population $E(k, t)$ as a function of time for $k = 0, \dots, 4$ as the global maximum of the fitness landscape in Fig. 2 lies at a Hamming distance 4 from the initial sequence. Since a line $E(k, t)$ intersects $E(k', t)$ at a time

$$T = \frac{k' - k}{F(k') - F(k)}, \tag{14}$$

and as $F(1) > F(0)$, the population of the sequence in shell 0 is overtaken by the one in shell 1. However to become a most populated genotype, it is not sufficient to have a record fitness. As Fig. 3 shows, although $F(2) > F(1)$, the population $E(2, t)$ overtakes $E(1, t)$ later than $E(3, t)$, losing the evolutionary race. Thus, only those set of sequences become leader that manage to overtake the current leader in the least time amongst all contenders. Finally, the globally fittest sequence $E(4, t)$ catches up with $E(3, t)$ and the population localizes at the global peak [35,38].

The shell model defined by (13) describes a population growing linearly in time with a slope equal to the fitness $F(k)$ chosen from the distribution $p_k(F)$. A simpler version of this model in which the fitness $F(k)$ is assumed to be independent and identically distributed (i.i.d.) variables with distribution $p(F)$ have also been considered [35,38]. Several properties of this model have been recently calculated via a mapping to a first-passage problem [44] and considering it as a system of

hard-core particles undergoing elastic collisions [27]. In particular, it has been shown that the average number of jumps grows as $\beta \ln L$ where the prefactor $\beta < 1$ and depends on the choice of $p(F)$. In both of these approaches, the initial condition (the intercept) of $E(k,t)$ is not fixed and is a uniformly distributed random number on the real line. In this paper, we present a way to calculate average number $\mathcal{J}$ of changes in k^* which respects the discreteness of the underlying genotypic space. We will perform the calculations for the shell model for which the fitness is nonidentically distributed, although our method is readily applicable to the i.i.d. model also. However, we mention that the prefactor β calculated using our method turns out to be the same as in the analysis of [27,44].

A. General formulas

To calculate the statistics of jumps, we need two basic distributions: (i) the probability $P_k(F,t)$ that the most populated sequence in shell k with fitness F is the leader at time t, and (ii) the probability $W_{k',k}(F,t)dt$ with which this sequence is overtaken by the most populated sequence in shell $k'(>k)$ between time t and $t+dt$. The distribution $P_k(F,t)$ requires that the population $E(j,t) < E(k,t)$, $j \neq k$ which implies that the fitness $F(j) < F+(j-k)/t$. Since the fitnesses of the most populated sequence in each shell are independent random variables, we have

$$P_k(F,t) = p_k(F) \prod_{\substack{j=0 \\ j \neq k}}^{L} \int_{F_{\min}}^{F+(j-k)/t} dF' p_j(F'). \qquad (15)$$

To find $W_{k',k}(F,t)$, we need to determine the fitness of the most populated sequence in shell $k'(>k)$ which can contribute to the overtaking event. Let us denote the location of the leader at time t by $E(t)$,

$$E(t) = -k + Ft.$$

Then the sequence in shell k' can overtake the kth one at t if the fitness $F(k') = (E(t)+k')/t$. Similarly, the sequence in the kth shell can be overtaken at $t+dt$, $dt/t \to 0$ if

$$F(k') = \frac{E(t+dt)+k'}{t+dt} = F + \frac{k'-k}{t} - \frac{k'-k}{t^2}dt + \mathcal{O}(dt^2).$$

Thus the probability that a sequence in shell k is overtaken by a sequence in shell k' between time t and $t+dt$ is given by

$$\begin{aligned} W_{k',k}(F,t)dt &= \int_{F+(k'-k)/t-[(k'-k)/t^2]dt}^{F+(k'-k)/t} dF p_k(F) \\ &\approx \frac{k'-k}{t^2} p_{k'}\left(F + \frac{k'-k}{t}\right)dt, \quad k' > k. \end{aligned} \qquad (16)$$

Finally, using the distributions defined in (15) and (16), we can write the probability $\mathcal{P}_{k',k}(t)$ that the most populated sequence in shell k' overtakes the one in shell k at time t as

$$\mathcal{P}_{k',k}(t) = \int_{F_{\min}}^{F_{\max}} dF W_{k',k}(F,t) P_k(F,t), \qquad (17)$$

where $F_{\max}$ is the upper support of the distribution $p(F)$.

Depending on the quantity of interest, one can either integrate over time or sum over a space variable in $\mathcal{P}_{k',k}(t)$. Often the experimental data such as a morphological feature [4] or average fitness [5] are plotted as a function of time and can be used to find the number of jumps in time. Therefore it is useful to consider the distribution $\mathcal{J}(t)$ of a jump to occur at time t which can be found by summing $\mathcal{P}_{k',k}(t)$ over k and k',

$$\mathcal{J}(t) = \sum_{k=0}^{L} \sum_{k'=k}^{L} \mathcal{P}_{k',k}(t). \qquad (18)$$

The relationship between the overtaken and the overtaking sequence can be deduced by computing the jump distribution $\mathcal{J}_k$ that the most populated sequence in shell k is a jump given by

$$\mathcal{J}_k = \sum_{k'=k}^{L} \int_0^\infty dt \mathcal{P}_{k',k}(t). \qquad (19)$$

The average number $\mathcal{J}$ of jumps can be obtained by either integrating $\mathcal{J}(t)$ over time or summing $\mathcal{J}_k$ over k.

In the previous works on shell models [35,37,38], several properties of the jumps have been studied numerically when the fitness F is distributed according to an exponential or Gaussian distribution. In the following subsections, we will use the expressions derived above for the exponential case which lends itself to a detailed analysis and then give some results for fitness distributions decaying faster than an exponential.

B. Exponentially distributed fitness

We consider $p(F) = e^{-F}$ for which the largest fitness in shell k is distributed according to

$$p_k(F) = \alpha_k e^{-F}(1 - e^{-F})^{\alpha_k - 1}. \qquad (20)$$

Using this in (15), we obtain the distribution that the leader with fitness F is in shell k,

$$\begin{aligned} P_k(F,t) &= \alpha_k \left(\frac{e^{-F}}{1-e^{-F}}\right) \prod_{j=0}^{L} (1 - e^{-F+(k-j)/t})^{\alpha_j} \\ &\approx \alpha_k e^{-F} e^{-(1+e^{-1/t})^L e^{-F+k/t}}, \end{aligned} \qquad (21)$$

where the last expression is obtained by exponentiating the product and keeping only the leading order terms in the expansion. Similarly, the overtaking rate (16) can be written as

$$\begin{aligned} W_{k',k}(F,t) &= \alpha_{k'}\left(\frac{k'-k}{t^2}\right) e^{-F+(k-k')/t}(1 - e^{-F+(k-k')/t})^{\alpha_{k'}-1} \\ &\approx \alpha_{k'}\left(\frac{k'-k}{t^2}\right) e^{-F+(k-k')/t} e^{-\alpha_{k'} e^{-F+(k-k')/t}}. \end{aligned} \qquad (22)$$

Then the probability that the population in the k'th shell exceeds the population in the kth shell at time t is given by
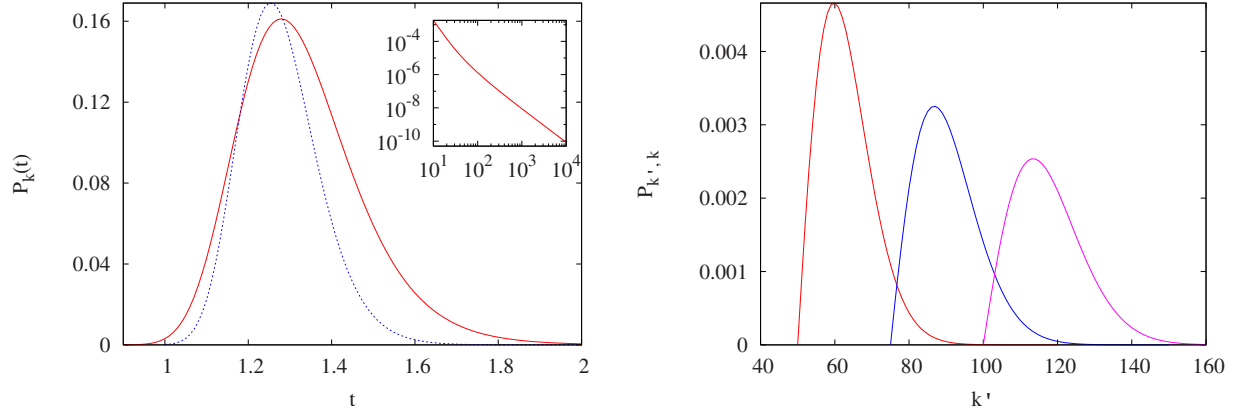
FIG. 4. (Color online) Jump statistics for $p(F)=e^{-F}$. Left: Distribution $\mathcal{P}_k(t)$ given by (25) for fixed $k/L=0.3$ in support of (28) for L =500 (solid) and 1000 (dotted). The inset shows the $1/t^2$ dependence of $\mathcal{P}_k(t)$ for $k=450$, $L=1000$. Right: Distribution $\mathcal{P}_{k',k}$ given by (32) for $k=50$, 75 and 100 (left to right) and $L=1000$.

$$\mathcal{P}_{k',k}(t) = \alpha_k \alpha_{k'} e^{(k-k')/t} \left(\frac{k'-k}{t^2} \right) \int_0^1 dz \, z \, e^{-z[\alpha_{k'} e^{-k'/t} + (1+e^{-1/t})L] e^{k/t}}.$$

(23)

Neglecting the first term in the exponent of the exponential in the integrand and carrying out the integral for large L, we finally obtain

$$\mathcal{P}_{k',k}(t) \approx \frac{\alpha_{k'} e^{-k'/t}}{(1+e^{-1/t})^L} \frac{\alpha_k e^{-k/t}}{(1+e^{-1/t})^L} \left(\frac{k'-k}{t^2} \right).$$

(24)

We will now use this expression to calculate jump statistics.

Temporal behavior. Let us first consider the distribution $\mathcal{P}_k(t)$ of a jump to occur in the kth shell at time t. Summing $\mathcal{P}_{k',k}(t)$ over k', we obtain [45]

$$\mathcal{P}_k(t) = \frac{\alpha_k e^{-k/t}}{t^2 (1+e^{-1/t})^{2L}} \sum_{k'=k}^L (k'-k) \alpha_{k'} e^{-k'/t}$$

$$= \frac{\alpha_k e^{-k/t}}{t^2(1+e^{-1/t})^{2L}} \frac{\Gamma(L+1)e^{-(k+1)/t}}{\Gamma(L-k)\Gamma(k+2)}$$

$$\times {}_2F_1(2, k+1-L; k+2; -e^{-1/t})$$

(25)

where $\Gamma(n)$ is the gamma function and ${}_2F_1(a,b;c;z)$ is the hypergeometric function. We point out that the distribution $\mathcal{P}_k(t)$ gives the probability that $E(k,t)$ is *overtaken* at t and hence differs from the distribution that $E(k,t)$ is the largest at time t considered in Ref. [38]. The function $\mathcal{P}_k(t)$ is plotted as a function of time for various values of k in Fig. 4. To gain some insight into the behavior of this distribution, we calculate the above sum using saddle point approximation. Using the Stirling's formula for binomial coefficient

$$\binom{L}{k} \approx \sqrt{\frac{L}{2\pi k(L-k)}} \frac{L^L}{k^k (L-k)^{L-k}}$$

(26)

in (25) for $\alpha_{k'}$ and approximating the sum over k' by an integral, we have

$$\int_k^L dk' (k'-k) e^{-f(k')}$$

$$\approx \frac{\sqrt{\pi} e^{-k_0'/t}}{f''(k_0')} \binom{L}{k_0'} \left(\Delta(k)\{\mathrm{erf}[\Delta(k)] - \mathrm{erf}[\Delta(L)]\} \right.$$

$$\left. + \sqrt{\frac{1}{\pi}} (e^{-\Delta^2(k)} - e^{-\Delta^2(L)}) \right),$$

where we have estimated the integral using the saddle point method. In the above expression, $f(k') = (k'/t) - \ln \alpha_{k'}$, $f''(k_0')$ is the second derivative of $f(k')$ evaluated at the minimum k_0' of the function $f(k')$ and the deviation $\Delta(k') = (k' - k_0') \sqrt{f''(k_0')/2}$. Explicitly,

$$k_0' = \frac{L}{1 + e^{1/t}},$$

$$f''(k_0') = \frac{L}{k_0'(L-k_0')} = \frac{(1+e^{1/t})(1+e^{-1/t})}{L},$$

$$\Delta(k) = (k - k_0') \sqrt{\frac{f''(k_0')}{2}}.$$

Approximating the factor $\alpha_k e^{-k/t}$ in (25) also by a Gaussian in a manner similar to above, we finally have

$$\mathcal{P}_k(t) = \frac{e^{-\Delta^2(k)}}{2\sqrt{\pi} t^2} \left(\Delta(k)\{\mathrm{erf}[\Delta(k)] - \mathrm{erf}[\Delta(L)]\} \right.$$

$$\left. + \sqrt{\frac{1}{\pi}} (e^{-\Delta^2(k)} - e^{-\Delta^2(L)}) \right).$$

(27)

Using this expression, it is easy to obtain the typical shell location of the sequence overtaken at time t by integrating $k\mathcal{P}_k(t)$ over k. Since $\mathcal{P}_k(t)$ is normalized to $\mathcal{J}(t)$, we find that given an overtaking event occurs at t, the average location of the overtaken sequence scales as k_0' with a standard deviation of order $1/\sqrt{f''(k_0')}$ about it. This length scale can also be

expressed in terms of time for fixed k as the distribution $\mathcal{P}_k(t)$ is maximized at $\Delta(k)=0$ or $k_0'(t)=k$. Thus for large L, the time $T_{\mathrm{m.p.}}$ at which $\mathcal{P}_k(t)$ is *most probable* is given by

$$T_{\mathrm{m.p.}} = \left[\ln\left(\frac{L-k}{k} \right) \right]^{-1}. \tag{28}$$

This means that a sequence with $k \sim \mathcal{O}(L)$ is most likely to be overtaken in a time of order unity. This fact is also expressed in Fig. 4 which shows the distribution $\mathcal{P}_k(t)$ at fixed k/L as a function of time. The time scale $T_{\mathrm{m.p.}}$ can be understood by a simple argument which estimates the intersection time given by (10). This argument is analogous to that used in Refs. [20,35,38] where the fitness difference in the denominator of (10) is given by the typical value of the fitness gap which probes the rare events. As we are interested in the most likely events, the denominator is approximated by the difference in the average value of the largest fitness in shell k' and k. From (20), we see that the average largest fitness goes as $\ln \alpha_k$. Since typically the most populated sequence in shell k is overtaken by a sequence located $\mathcal{O}(\sqrt{k})$ distance away [38] (also see below), the numerator of (10) scales as $\sqrt{k}$ and the denominator $\ln(\alpha_k'/\alpha_k)$ on using the Stirling's formula turns out to be $\sqrt{k}\ln((L-k)/k)$ thus leading to (28).

Although the most probable value of the overtaking time is of order one, the *average* overtaking time is infinite due to the fat tail of the distribution $\mathcal{P}_k(t)$. For $t \gg 1$, we can approximate $\Delta(k)$ by $(2k-L)/\sqrt{2L}$ and using the asymptotic expansion of error function for large argument [46],

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x dy\, e^{-y^2} = 1 - \frac{e^{-x^2}}{\sqrt{\pi}x}\left(1 - \frac{1}{2x^2} + \cdots \right), \tag{29}$$

we obtain

$$\mathcal{P}_k(t) \approx \sqrt{\frac{L}{2\pi}} \frac{\epsilon}{t^2} e^{-L\epsilon^2/2}, \quad \epsilon = 1 - \frac{2k}{L} > 0. \tag{30}$$

Due to the t^{-2} behavior at large times shown in the inset of Fig. 4, the mean time diverges for any L and $k < L/2$ (see below). The tail of this distribution is exponentially suppressed in L for finite ϵ but goes as $\sqrt{L}\epsilon/t^2$ for k close to $L/2$. The late time behavior above is also obtainable from (10) by a simple change of variables [35,38].

After performing the integral over k in (27), we obtain

$$\mathcal{J}(t) = \sqrt{\frac{L}{4\pi}} \frac{1}{t^2} \mathrm{sech}\left(\frac{1}{2t} \right). \tag{31}$$

Integrating $\mathcal{J}(t)$ over time from 0 to infinity, we find that the average number of jumps grows as $\sqrt{L\pi}/2$.

Spatial behavior. One can also find the probability $\mathcal{J}(k)$ that the most populated sequence in the kth shell is a jump. As we are not interested in temporal distribution, the integral over time in (24) can be carried out to give the probability that the sequence in shell k is overtaken by that in shell k',

$$\mathcal{P}_{k',k} = \int_0^\infty dt \mathcal{P}_{k',k}(t)$$

$$= \frac{k'-k}{k'+k}\binom{L}{k}\binom{L}{k'} {}_2F_1(k+k',2L;k+k'+1;-1). \tag{32}$$

This distribution is shown for some representative parameters in Fig. 4. Approximating the integrand $\mathcal{P}_{k',k}(t)$ in the above equation by a Gaussian centered about inverse time $t^{-1} = -\ln((k'+k)/(2L-k'-k))$ and carrying out the integral, we obtain

$$\mathcal{P}_{k',k} \approx \frac{L(k'-k)}{(k+k')(2L-k'-k)}\binom{L}{k}\binom{L}{k'}\binom{2L}{k'+k}^{-1}$$
$$\times \left\{ 1 + \mathrm{erf}\left[\sqrt{\frac{(k+k')(2L-k-k')}{4L}} \right.\right.$$
$$\left.\left. \times \ln\left(\frac{2L-k'-k}{k'+k} \right) \right] \right\}. \tag{33}$$

The argument of the error function changes sign when $k'+k=L$. For $k'+k>L$, the argument is negative and of order $\sqrt{L} \gg 1$. Using (29), we find that the last factor in (33) is exponentially small in L for $k'>L/2, k<k'$. Thus the probability that the overtaking sequence k' lies beyond the shell $L/2$ is negligible. This is understandable as the globally fittest sequence is typically located in the shell $k=L/2$ [38]. For $k,k'<L/2$, the error function in (33) can be approximated by unity, and the probability distribution $\mathcal{P}_{k',k}$ can be further simplified to give

$$\mathcal{P}_{k',k} \approx \sqrt{\frac{L}{\pi k(L-k)}}\left(\frac{k'-k}{2k} \right) e^{-L(k'-k)^2/4k(L-k)},$$

$$k < k' < L/2 \tag{34}$$

where we have used the Gaussian approximation for the binomial coefficients. This form of the distribution implies that the overtaking sequence k' is located within $\mathcal{O}(\sqrt{k})$ distance of the overtaken sequence k. Thus the typical spacing between successive jumps for large k is roughly constant and goes as $\sqrt{L}$ as seen in the numerical simulations of Ref. [38]. The jump distribution $\mathcal{J}_k$ for a jump to occur in shell k is obtained by integrating over k' and we have

$$\mathcal{J}_k \approx \sqrt{\frac{L}{\pi k(L-k)}}\, \theta_H\left(\frac{L}{2} - k \right), \tag{35}$$

where θ_H is the Heaviside step function. Thus the distribution $\mathcal{J}_k$ decays as $k^{-1/2}$ for $k \ll L$ in accordance with the numerical results of Refs. [37,38]. Integrating the preceding equation over k, we find that the average number of jumps are given as $\sqrt{L\pi}/2$ in agreement with the previous calculation.

C. Gumbel-distributed shell fitness

We now consider $p(F)=Ae^{-F^\gamma}$, $\gamma>0$, $A=\gamma/\Gamma(1/\gamma)$ for which the distribution $p_k(F)$ of the largest of α_k random vari-

ables for large L has the Gumbel distribution as the limiting form [47],

$$p_k(F) = \frac{1}{C_k} e^{-(F-B_k)/C_k} e^{-e^{-(F-B_k)/C_k}}, \quad (36)$$

where

$$B_k = \left[\ln\left(\frac{\alpha_k}{\Gamma(1/\gamma)}\right)\right]^{1/\gamma}, \quad C_k = \frac{1}{\gamma}\left[\ln\left(\frac{\alpha_k}{\Gamma(1/\gamma)}\right)\right]^{(1-\gamma)/\gamma}. \quad (37)$$

We will show that the tail of the distribution $\mathcal{P}_k(t)$ decays as $1/t^2$ and the average number of jumps scales as $\sqrt{L}$ for any $\gamma > 0$.

The large time behavior of $\mathcal{P}_k(t)$ can be found by taking $t \to \infty$ limit in (15) and (16) except for the $1/t^2$ factor in rate $W_{k',k}$. We thus have

$$\mathcal{P}_k(t) \approx \sum_{k'=k}^{L} \frac{k'-k}{t^2} \frac{e^{B'_k/C'_k + B_k/C_k}}{C'_k C_k}$$
$$\times \int_0^\infty dF e^{-F/C_k} e^{-F/C'_k} e^{-\Sigma_{j=0}^L e^{B_j/C_j} e^{-F/C_j}}, \quad t \gg T_{\text{m.p.}} \quad (38)$$

after using the approximations similar to those used in arriving at (24). The sum under the integral sign can be computed by saddle point approximation so that the integral (up to scale factors) is writable as

$$\int_0^\infty dz z^{\left(\frac{\ln \alpha'_k}{L \ln 2}\right)^{(\gamma-1)/\gamma} + \left(\frac{\ln \alpha_k}{L \ln 2}\right)^{(\gamma-1)/\gamma} - 1} e^{-z},$$

where we have neglected the logarithm of $\Gamma(1/\gamma)$ in B_k and C_k for large L. Since we are interested in large times when $k \to L/2$, to leading order in $\epsilon = 1 - (2k/L)$, we find $\ln \alpha_k / L \ln 2 \approx 1$, thus simplifying the expression for $\mathcal{P}_k(t)$ to yield

$$\mathcal{P}_k(t) = \left(\frac{\gamma}{t}\right)^2 \left(\frac{\pi L}{2}\right)^{\gamma-1} (L \ln 2)^{(\gamma-1)/\gamma} \left(\frac{\alpha_k}{2^L}\right)^\gamma \sum_{k'=k}^{L} (k'-k)\left(\frac{\alpha'_k}{2^L}\right)^\gamma$$
$$= \sqrt{\frac{L}{2\pi\gamma}} \frac{\gamma^2 \epsilon}{t^2} (L \ln 2)^{(\gamma-1)/\gamma}, \quad t \gg T_{\text{m.p.}} \quad (39)$$

where the last expression is obtained by estimating the sum over k' using Gaussian approximation. For small ϵ ($\sim 1/\sqrt{L}$) and $\gamma = 1$, this expression reduces to (30) for exponential distribution. For $\gamma \neq 1$, the L dependence in the preceding equation consists of two factors, the first one arising because the typical separation $k'-k \sim \mathcal{O}(\sqrt{L})$ for large k [35,38] and the last factor is due to $1/C_k$, $k=L/2$ left after scaling the fitness F in (38) by $C_{k'}$.

In order to find the average number of jumps, we first need to calculate $T_{\text{m.p.}}$ for arbitrary γ. As argued in the last subsection, $T_{\text{m.p.}}^{-1}$ is given as the derivative (with respect to k) of the average shell fitness. From the scaling form (36), we see that the average shell fitness is proportional to B_k. For k

of order L, its derivative grows as C_k, $k \sim L$. Then integrating $\mathcal{P}_k(t)$ over k and t, we find that the average number of jumps scale as $\sqrt{L}$ for all $\gamma > 0$. This result is also consistent with the simulations for the Gaussian distribution in Ref. [37].

V. JUMP DISTRIBUTION FOR FINITE POPULATIONS

We will now consider the dynamics of a population of N individuals evolving according to the Wright-Fisher dynamics described in Sec. II. Unlike the infinite population, a population of size N initially localized at sequence $\sigma^{(0)}$ can spread up to a finite distance. This is since the typical fraction of the population at a sequence σ in one generation is given as $\mu^{d(\sigma,\sigma^{(0)})}$ but as this fraction is bounded below by $1/N$, it follows that the mutational spread $d_{\text{eff}} = \ln N/|\ln \mu|$ for a finite population. Thus while all the mutants are available in one generation for quasispecies [see (5)], only a finite number is present at any time in real populations. If a more fit sequence is available within this effective distance, the population behaves like a quasispecies and evolves deterministically. This is possible for $d_{\text{eff}} \gtrsim 1$ and at short times [20]. However, at long times, any finite population can get trapped at a local peak if a more fit sequence lies farther out than d_{eff}. In such an event, the population escapes the local peak via the process of stochastic tunneling which takes a time given by [39,40]

$$\Delta T = \left[N\mu^2 L\left(\frac{W_{\text{loc,f}} - W_{\text{loc,i}}}{W_{\text{loc,f}} W_{\text{loc,i}}}\right)\right]^{-1} \quad (40)$$

for $d_{\text{eff}} = 1$, where $W_{\text{loc},[i,f]}$ refers to the fitness of the initial and final local peak separated by two mutations. During this time, most of the population stays at the local peak with fitness $W_{\text{loc,i}}$ but a few less-fit mutants are produced by single mutations. When some of these mutants further acquire a favorable mutation, then the whole population quickly jumps to the next local peak with higher fitness $W_{\text{loc,f}}$. The physical process involved when a jump occurs in a finite population is thus different from that in the quasispecies case. In the latter case, each local peak is already populated albeit with a small frequency, and a jump occurs when the population at a more fit sequence overtakes the current leader.

We are interested in the jump distribution of large finite populations with d_{eff} close to one. At large times when the tunneling drives the dynamics, we expect the density $J(t)$ of jumps at t to scale as

$$J(t) \sim \frac{1}{\Delta T} \sim N\mu^2 L \frac{\Delta w(t)}{w_i^2(t)}, \quad (41)$$

where $w_i(t)$ is the typical fitness of the local peak visited at t separated by a better peak with fitness difference $\Delta w(t)$. We expect Δw to decrease and w_i to increase with time as higher peaks are explored. However, in the absence of an argument for these time dependences, we present our preliminary numerical results here. In the shell model, one does not have to deal with the whole genotypic space consisting of 2^L sites and it suffices to work with the L shells, thus reducing the computational effort enormously [35]. However, such a rota-
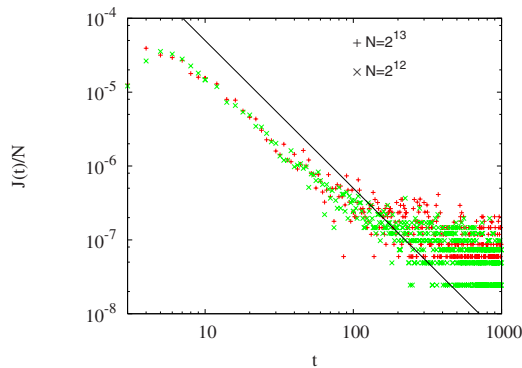
FIG. 5. (Color online) Scaled jump distribution $J(t)/N$ for a finite population with N individuals to show the $1/t^2$ dependence. Here $L=6$, $\mu=0.0004$, $p(W)=W^{-2}$ and the data have been averaged over 500 fitness landscapes with 10 histories each.

tional symmetry is not present for the finite population problem so we are able to handle only small values of L. Our numerical results for large finite populations and small L are shown in Fig. 5 for exponentially distributed log fitness F or $p(W)=W^{-2}$, as in the previous sections. The data in Fig. 5 are averaged over several histories as the evolutionary trajectories are not deterministic for finite populations [20]. For fixed μ and L, we find that at long times,

$$J(t) \sim \frac{N}{t^2}, \qquad (42)$$

which decays the same way as in the quasispecies model,

$$\mathcal{J}(t) \sim \left(\frac{\ln \mu}{t} \right)^2 \sqrt{L}, \quad t \gg 1, \qquad (43)$$

where we have reinstated the μ dependence.

VI. CONCLUSIONS

In this paper, we discussed the evolution of asexual populations on rugged fitness landscapes with many local optima separated by valleys. We focused on the statistical properties of the most populated genotype which changes as the population locates better peaks in the fitness landscape. These properties were calculated exactly within a shell model which was derived systematically from the Eigen's quasispecies model for infinite populations. We showed that the expression for the population frequency within the shell model approximates the quasispecies solution well for highly fit sequences and at short times only. However, the two models are equivalent as regards the statistics of the most populated genotype. We computed the average number of jumps in the shell model and found that it grows as $\sqrt{L}$, L being the sequence length, for fitness distributions decaying as exponential or faster. The jump distribution in time was shown to decay as t^{-2}. This dependence is also seen numerically for the finite population but a satisfactory explanation for this case is presently missing. A more detailed analysis of the finite population properties will be taken up in the future.

[1] W. Fitch and F. Ayala, Proc. Natl. Acad. Sci. U.S.A. **91**, 6717 (1994).

[2] E. Mayr, *Animal Species and Evolution* (Harvard University Press, Cambridge, 1963).

[3] N. Eldredge and S. Gould, in *Models in Paleobiology*, edited by T. Schopf (Freeman, Cooper and Co., San Francisco, 1972), pp. 82–115.

[4] M. Benton and P. Pearson, Trends Ecol. Evol. **16**, 405 (2001).

[5] S. F. Elena and R. E. Lenski, Nat. Rev. Genet. **4**, 457 (2003).

[6] I. Novella, E. Duarte, S. Elena, A. Moya, E. Domingo, and J. Holland, Proc. Natl. Acad. Sci. U.S.A. **92**, 5841 (1995).

[7] R. E. Lenski and M. Travisano, Proc. Natl. Acad. Sci. U.S.A. **91**, 6808 (1994).

[8] M. Imhof and C. Schlotterer, Proc. Natl. Acad. Sci. U.S.A. **98**, 1113 (2001).

[9] C. L. Burch and L. Chao, Genetics **151**, 921 (1999).

[10] S. Wright, in *Proceedings of the Sixth International Congress of Genetics*, edited by D. F. Jones (Brooklyn Botanic Garden, Menasha, Wisconsin, 1932), pp. 356–366.

[11] P. Stadler, in *Biological Evolution and Statistical Physics*, edited by M. Lässig and A. Valleriani (Springer, Berlin, 2002), p. 183.

[12] S. Gavrilets, *Fitness Landscapes and the Origin of Species* (Princeton University Press, New Jersey, 2004).

[13] L. S. Tsimring, H. Levine, and D. A. Kessler, Phys. Rev. Lett. **76**, 4440 (1996).

[14] S. A. Kauffman, *The Origins of Order* (Oxford University Press, New York, 1993).

[15] R. Korona, C. H. Nakatsu, L. J. Forney, and R. E. Lenski, Proc. Natl. Acad. Sci. U.S.A. **91**, 9037 (1994).

[16] C. L. Burch and L. Chao, Nature (London) **406**, 625 (2000).

[17] F. Poelwijk, D. Kivet, D. Weinreich, and S. Tans, Nature (London) **445**, 383 (2007).

[18] G. Fernandez, B. Clotet, and M. Martinez, J. Virol. **81**, 2485 (2007).

[19] M. C. Whitlock, P. C. Phillips, F. B.-G. Moore, and S. J. Tonsor, Annu. Rev. Ecol. Syst. **26**, 601 (1995).

[20] K. Jain and J. Krug, Genetics **175**, 1275 (2007).

[21] B. Derrida, Phys. Rev. B **24**, 2613 (1981).

[22] D. Sherrington and S. Kirkpatrick, Phys. Rev. Lett. **35**, 1792 (1975).

[23] P. Sibani and J. Dall, Europhys. Lett. **64**, 8 (2003).

[24] J. Krug and T. Halpin-Healy, J. Phys. I **3**, 2179 (1993).

[25] P. L. Krapivsky and S. Redner, Phys. Rev. Lett. **89**, 258703 (2002).

[26] E. Ben-Naim and P. Krapivsky, Europhys. Lett. **65**, 151 (2004).

[27] I. Bena and S. N. Majumdar, Phys. Rev. E **75**, 051103 (2007).

[28] S. Sabhapandit, J. Stat. Mech.: Theory Exp. (2007) L05002.

[29] M. Eigen, Naturwiss. **58**, 465 (1971).

[30] K. Jain and J. Krug, in *Structural Approaches to Sequence Evolution: Molecules, Networks and Populations*, edited by U. Bastolla, M. Porto, H. Roman, and M. Vendruscolo (Springer, Berlin, 2007), pp. 299–340.

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

[32] C. Kamp and S. Bornholdt, Phys. Rev. Lett. **88**, 068104 (2002).

[33] D. B. Saakian, E. Muñoz, C.-K. Hu, and M. W. Deem, Phys. Rev. E **73**, 041913 (2006).

[34] J.-M. Park and M. W. Deem, Phys. Rev. Lett. **98**, 058101 (2007).

[35] J. Krug and C. Karl, Physica A **318**, 137 (2003).

[36] Y.-C. Zhang, Phys. Rev. E **55**, R3817 (1997).

[37] J. Krug and K. Jain, Physica A **358**, 1 (2005).

[38] K. Jain and J. Krug, J. Stat. Mech.: Theory Exp. (2005) P04008.

[39] Y. Iwasa, F. Michor, and M. A. Nowak, Genetics **166**, 1571 (2004).

[40] D. M. Weinreich and L. Chao, Evolution (Lawrence, Kans.) **59**, 1175 (2005).

[41] H. Flyvbjerg and B. Lautrup, Phys. Rev. A **46**, 6714 (1992).

[42] D. Posada, K. Crandall, and E. Holmes, Annu. Rev. Genet. **36**, 75 (2002).

[43] J. Drake, B. Charlesworth, D. Charlesworth, and J. Crow, Genetics **148**, 1667 (1998).

[44] C. Sire, S. Majumdar, and D. S. Dean, J. Stat. Mech.: Theory Exp. (2006) L07001.

[45] I. Gradshteyn and I. Ryzhik, *Table of Integrals, Series, and Products* (Academic, New York, 1980).

[46] G. Arfken, *Mathematical Methods for Physicists* (Academic, New York, 1985).

[47] H. A. David, *Order Statistics* (Wiley, New York, 1970).