

Statistics of an Age Structured Population with Two Competing Species: Analytic and Monte Carlo Studies

R.K.P. Zia^{1,2} and R.J. Astalos¹

¹ Center for Stochastic Processes in Science and Engineering,
Department of Physics,
Virginia Polytechnic Institute and State University, Blacksburg, VA,
24061-0435, USA

² Fachbereich Physik, Universität - Gesamthochschule Essen,
D-45117 Essen, Federal Republic of Germany

Abstract. Using Monte Carlo simulations and analytic methods, we study the effects of competition and mutation between species with age structure. For definiteness, simulations are based on the reproductive and age structures of the Penna-Desai model. In general, even mean-field (deterministic) type of analytic results are not available. However, by focusing on just *two* species, we find successful analytic approaches to the observed *stochastic* aspects of the population dynamics.

1 Introduction

Within the statistical physics community, there has been considerable renewed interest in the venerable problem of population dynamics [1–3], in the context of bit-string models [4–6]. Introduced by Penna [4] as a model for the aging process, a system consists of individuals associated with a “genome,” which is an ordered string of good and bad “genes” represented by bits (0 and 1, respectively). Evolving discretely in time, each individual ages by a “year” and the next “gene” along the sequence is expressed. The number of bad “genes” (deleterious mutations) *accumulated* up to that time dictates the probability of survival and reproduction. When (*asexual*) reproduction occurs, the offspring inherit the same genome, unless mutations are allowed. Of course, being a newborn, that individual would have *zero* accumulated deleterious mutations. Only in subsequent years does it suffer from such accumulations. In most studies, competition for resources is introduced in the simplest manner, through a Verhulst factor [1], so that the total population typically reaches a non-trivial steady state. Of interest are the effects of both competition and mutation on the nature of the final state, as well as various time-dependent phenomena.

In most of the recent literature, the length of the bit-strings used were 32 or 64, so that the total number of possible “species” is 2^{32} or 2^{64} . If mutations are allowed, all species may, in principle, be present in the population. As

a non-equilibrium statistical mechanics system, even the steady state properties, never mind time-dependent aspects, are typically intractable analytically. If, on the other hand, no mutations are allowed, the system resembles a thermodynamic system at zero temperature, so that its final state depends on initial conditions, i.e., the specific gene pool at $t = 0$. In order to make some progress into a more detailed understanding of the combined effects of competition and mutations, we choose an alternative route, namely, restricting our considerations to only *two* genomes. This study should be viewed as a small, initial step towards the goal of predicting properties of systems with 2^{32} (or more) species. Our approach involves both Monte Carlo simulations of populations (within the paradigm of Penna models) and various analytic approaches (with less model dependence). If the “fitness” of the two species are quite distinct, then a deterministic, “mean-field” description is adequate: the fitter genome dominates in late times (for reasonable μ 's). However, if the genomes are equally fit, this approach fails miserably, by predicting that f , the fraction of the population being one of the species, is *time independent*. A naive guess for our *stochastic* process might be a random walk in the interval $f \in [0, 1]$, with absorbing or “sticky” boundaries for $\mu \equiv 0$ or $\ll 1$. In this paper, we show that the problem is more complex, displaying interesting features beyond simple expectations. Using simulations, we measure $P(f, t; \mu)$, the probability of the population at time t to consist of a fraction f of one species (and $1 - f$ of the other), for a fixed mutation rate μ . The appropriate description is, instead of a simple random walk, Fisher-Wright diffusion [2]. For the readers' convenience, we provide the details of the Penna-Desai model, on which our simulations are based. Two subsequent sections are devoted to the dynamics of two competing species, with and without mutation. Monte Carlo results, as well as various levels of analytic approaches applicable beyond the Penna-Desai model, are presented. A summary is given in the last section, as well as speculations on interesting generalizations of this simple model.

2 The bit-string model for age-structured populations

Although our aim is to arrive at model-independent conclusions, we must perform simulations with specific models. In addition, it is easier to focus our theoretical discussions around a particular example. For this purpose, we choose a model recently introduced by Penna [4], with extensions by Desai, et. al. [6]. For completeness, we describe this model in detail here.

At a given time t (which are discrete steps labeled as “years”), our population consists of $N(t)$ “individuals.” From one year to the next, all individuals age by a year, while some die and some are born. At some initial time ($t = 0$), the population may consist of only “babies” or a distribution of ages. Each individual is born with a genome g , represented by a string of L bits, i.e., $g = \{\sigma_i\}$, $i = 1, \dots, L$ with $\sigma = 0$ or 1 . Thus, the name, “bit-string model,”

was coined. In the following, we will refer to 0(1) as a “good” (“bad”) bit, or “gene”. Alternatively, a 1 bit can be regarded as a deleterious mutation in the genome. The central theme of Penna’s model of aging is that, as an individual ages from one year to the next, another bit in the string is “expressed.” That individual’s survivability and reproductivity then depends on the number of deleterious mutations *accumulated* up to that point. In other words, for each individual (of age α), we should focus on $b_g(\alpha) \equiv \sum_{i=1}^{\alpha} \sigma_i$, which controls its “fitness.” Given a specific set of rules for survival and reproduction, the dynamics of the population is well defined.

In the original Penna model [4], the reproductive rules are given by three parameters: R, B , and μ . The first represents age of “puberty,” in that individuals with $\alpha \leq R$ do not reproduce, regardless of g . The second models “menopause,” in that individuals with B or more accumulated deleterious mutations ($b_g(\alpha) \geq B$) stop reproducing. Thus, the age of menopause is α_M , which is defined through $b_g(\alpha_M) = B$. Clearly, this age depends on g ; but, for simplicity, we drop this subscript. Finally, μ is a measure of mutations, i.e., the genome of an offspring is, apart from randomly chosen μ bits, an exact copy of its parent. Since no other individual is involved in this activity, Penna’s model is associated with *asexual* reproduction.

Actually, there are other “hidden” parameters, in that every reproducing adult gives birth to a *single* “baby”. The generalizations by Desai [6] are two-fold. First, not every individual of reproductive age gives birth. Instead, the probability of reproducing, w , is a function of the number of deleterious mutations suffered, so that w depends on (g, α) through $b_g(\alpha)$. Second, if reproduction occurs, F babies are born, with F being also (g, α) dependent, through b . (Known as the fecundity, F is just the litter size.) Specifically,

$$w_g(\alpha) = e^{-b_g(\alpha)/B} \quad \text{and} \quad F_g(\alpha) = F_0 [1 - be^{b-B}/B] , \quad (1)$$

where F_0 is the maximum litter size. Thus, the average *productivity* of an adult of age α (in that year) is just the product:

$$p(\alpha) = w(\alpha)F(\alpha) , \quad (2)$$

where the index g has been suppressed for clarity. Recalling the condition of puberty, we set $p(\alpha) \equiv 0$ for $\alpha \leq R$. For the Penna model, we simply have $p(\alpha) = 1$ for $R < \alpha \leq \alpha_M$. In a mean-field type treatment, it is clear that only the functions $p(\alpha)$ play a role.

Next, we consider survivability. First, since there are only L bits in the string, no one survives beyond L years. Second, a parameter, T , is introduced to model the “terminal” number of accumulated deleterious mutations allowed, in that individuals with $b = T$ do not survive another year. Thus, we define the age of death, D_g , via

$$b_g(D_g) = T , \quad (3)$$

which shows that, when T is a genome independent parameter in the model, D_g depends on g through this equation. Finally, to model the carrying capacity of the environment, a Verhulst factor [1] is introduced to prevent the possibility of unlimited growth. In the simplest scenario, V depends only on the *total population*, \mathcal{N} , and represents the most elementary form of competition. It enters the dynamics in the form of an individual’s probability of surviving until the next breeding season. Thus, $V \in [0, 1]$ and should include a parameter which controls the population size. Further, the most reasonable choices for V would be monotonically decreasing functions of \mathcal{N} . A popular form is $V(\mathcal{N}) = 1 - \mathcal{N}/N_{\max}$ (0 for $\mathcal{N} > N_{\max}$), where N_{\max} is the maximum population the environment can support [4,6]. In our simulations, we prefer the form

$$V = \exp(-\mathcal{N}/\mathcal{N}_0) , \quad (4)$$

where \mathcal{N}_0 acts as a control parameter. With no *absolute* maximum population size, we may consider models with large fecundity, which could be susceptible to complete annihilation in a single year if subjected to $V = 1 - \mathcal{N}/N_{\max}$.

To summarize, we start with an initial population of a chosen distribution in age and genome (e.g., all babies, with randomly chosen g ’s) and update “yearly,” according to the following rules:

- compute \mathcal{N} and V from (4);
- let each individual survive with probability V ;
- let each survivor become a parent with probability $w_g(\alpha)$;
- let each parent reproduce $F_g(\alpha)$ babies;
- randomly choose μ bits in the genome of each baby and change them;
- increase the age of each adult by unity ($\alpha \rightarrow \alpha + 1$);
- eliminate those with $\alpha > D_g$ or L .

Note that $(L, T, R, B, F_0, \mu, \text{ and } \mathcal{N}_0)$ is a complete set of independent parameters, while the dependent ones are formed by Eqns (1,3,4). In practice, each individual is a bit-string, associated with a “pointer,” located at the α^{th} bit for its age.

Before continuing, let us introduce a simplifying modification to the mutation rule. Previous studies used the above rule of mutations and, for obvious reasons, found a proliferation of genomes, regardless of initial conditions. Since L was typically 32, the systems involve competition between roughly 10^{10} “species!” As a result, it is unclear what are the controlling factors for a particular species to be “successful” (in, say, the steady state). In particular, it is tempting to define “fitness” of a particular genome, g , by N_g^* , the *steady state population when no other species or mutation are present*. Certainly, according to mean-field predictions, a population evolving without mutations will eventually be completely dominated by the genome with the largest N^* . If small mutations are present, the steady state distribution is *not* necessarily controlled by this measure of “fitness” alone. Based on preliminary studies

of a simple $L = 3$ system with small μ , we find that steady state populations of genomes of the *same* N^* are *distinct* [7]. Instead, there is an additional, complex dependence on the “connectivity” between various species. For this reason, we restrict ourselves here to a study of the effects of mutations to just *two* genomes, by imposing the following rules.

- We choose two specific genomes to focus our investigations: g and g' .
- A newborn inherits the genome of the parent with probability $1 - \mu$.

In other words, with probability μ , a baby is given the genome of the *competing* species. While this restriction may not be ubiquitous in nature, we believe that such studies will give us some insights into the combined effects of competition and mutation on a population. Investigations of the more general cases are in progress, but they lie beyond the scope of this paper.

Given this set of rules, we turn our attention to properties of steady states and some time dependent phenomena. Clearly, the complete description of a population is given by $n_g(\alpha, t)$, the number of individuals with genome g , of age α , at time t . Of course, the total \mathcal{N} is just

$$\mathcal{N}(t) = \sum_{g, \alpha} n_g(\alpha, t) \quad (5)$$

and only plays a role for competition. However, due to the probabilistic rules, these are stochastic quantities. Thus, we may ask of their average values and, beyond that, fluctuations and correlations. Of course, the full information will be carried by $P(\{n_g(\alpha)\}; t)$, the probability, at time t , of finding the population with $n_g(\alpha)$ individuals of genome g and age α . In the single genome case (without mutation), this analysis has been performed [9]. Not surprisingly, for large \mathcal{N}_0 , with the population far from bifurcation points, the stationary distribution is a Gaussian. The first and second moments, corresponding to the mean values and the variances, respectively, agree well with simulation results. In this paper, we extend this study to the case of two genomes. First, we focus on competition *without* mutations. Then, mutations (in the restricted sense above) are included. Both simulations and analytic studies are presented. The extent of agreement between these two approaches allows us to conclude that even simple theoretical formulations capture the essence of these stochastic models of population dynamics.

3 Population dynamics of two competing species without mutation

It is well known, from both theoretical and field studies, that when two or more species compete for the same resources (in the simple fashion described above), only one tends to survive in the long run. This is known as the “principle of competitive exclusion” [8]. For species with very different “fitness,”

even the most simplistic mean-field theory succeeds in describing how the final state is reached. However, for competitive species, fluctuations dominate and mean-field approaches fail. Our interest lies in various levels of description beyond the mean-field.

3.1 The mean-field description

For the readers' convenience, we include here the well established mean-field approach, which neglects all stochastic aspects of the problem. It predicts that the species with the largest steady state population (N^*) will be the only survivor. Based on the rules specified above, the evolution equations for $n_g(\alpha, t)$, at this simplistic, deterministic level, can be written as:

$$n_g(0, t+1) = \sum_{\alpha=1}^L p_g(\alpha) V(\mathcal{N}(t)) n_g(\alpha, t) \quad (6)$$

$$\begin{aligned} n_g(\alpha+1, t+1) &= V(\mathcal{N}(t)) n_g(\alpha, t) \quad \text{for } \alpha \in [1, D_g - 1] \\ n_g(\alpha, t+1) &= 0 \quad \text{for } \alpha > D_g. \end{aligned} \quad (7)$$

Here, the first line accounts for the newborns (defined by $\alpha = 0$) while the second models the aging after surviving the competitive culling. The last line reflects the maximum age for g , so that the sum in the first equation effectively cuts off at $\alpha = D_g$. Note that, due to the dependence on \mathcal{N} (through Eqn 5), these equations are *nonlinear*. For the rest of the analysis, we will need only some general properties of p and V , so that the conclusions will be valid for a wide class of functional forms. In particular, we will focus our attention only on regions of parameter space which are associated with simple (period one) fixed points. The more complex issues of period doubling and chaos [10] should be addressed, but they lie beyond the scope of this paper.

First, despite the nonlinear nature of Eqns (6, 7), the equations for the steady state distribution, $n_g^*(\alpha)$, are *linear*:

$$n_g^*(0) = \sum_{\alpha} p_g(\alpha) V(\mathcal{N}^*) n_g^*(\alpha) \quad (8)$$

$$n_g^*(\alpha+1) = V(\mathcal{N}^*) n_g^*(\alpha) \quad \text{for } \alpha \in [1, D_g - 1], \quad (9)$$

since $V(\mathcal{N}^*)$ is just a number at the fixed point. If the system consists of only a *single genome*, which we label by G , then \mathcal{N}^* is just N_G^* . A simple recipe [11] for finding the non-trivial steady state begins with finding the *unique* solution to $\sum_{\alpha} p_G(\alpha) v^{\alpha+1} = 1$, with real $v \in [0, 1]$. (If the p 's are too small and the solution lies beyond $v = 1$, then the only stable steady state is complete extinction: $N_G^* = 0$.) Then, N_G^* is determined by the equation $v = V(N_G^*)$, while the entire age-distribution is given by

$$n_G^*(\alpha) = N_G^* v^{\alpha} (1 - v) / (1 - v^{1+D_G}) . \quad (10)$$

Furthermore, this fixed point is stable against small perturbations. If, on the other hand, the population consists of *many* species *and* if their reproductive rates lead to *non-degenerate* set of N^* 's, then it is straightforward to show that the only fixed point distributions consist of *single* genome cases. In addition, the one with the largest N^* is the most stable, which is the mathematical version of the Darwinian motto: “survival of the fittest.” In this case, where vying for a common resource accounts for the only competition, fitness is easily traced to reproductive rates, which are, in turn, monotonically related to N^* . That the steady population consists of this single genome is often referred to as the “Eve effect.” On the other hand, if two or more genomes share the same N^* (e.g., in the original Penna model), then there are fixed lines (or higher dimensional subspaces) which join the single genome fixed points. (For mathematical details leading to these conclusions, see, e.g., [7].)

Going beyond mean-field theory, it is easy to imagine that a noisy dynamics will change isolated fixed points into distributions well approximated by Gaussians. However, it is less clear what effects noise has on fixed lines (or higher dimensional subspaces). In the remainder of this section, we turn to the stochastic aspects of populations with just *two* competing species. We will be studying several cases, with different pairs of genomes, labeled by g and g' . To keep notations simple, given that we have only two genomes, we define

$$p_\alpha \equiv p_g(\alpha); q_\beta \equiv p_{g'}(\beta) \quad (11)$$

and

$$n_\alpha \equiv n_g(\alpha); m_\beta \equiv n_{g'}(\beta) \quad (12)$$

so that $N = \sum_\alpha n_\alpha$ and $M = \sum_\alpha m_\alpha$. The total population is just $\mathcal{N} = N + M$, and the fraction of \mathcal{N} being of genome g is defined by

$$f \equiv N / (N + M) . \quad (13)$$

3.2 Competition between very different genomes

As an “appetizer,” we consider genomes with very different N^* 's, so that the “fitter” species (one with larger N^*) should dominate the population rapidly. Specifically, we simulated the Penna model with generalized reproductive rates (1) and $L, T, R, B, F_0, \mathcal{N}_0 = 32, 5, 8, 5, 5, 20000$, respectively. The two genomes chosen are

$$g = 0000\ 0000\ 0000\ 0011\ 0111\ 0000\ 0000\ 0000 \quad (14)$$

$$g' = 0000\ 0000\ 0010\ 0000\ 1110\ 0000\ 0100\ 0000 \quad (15)$$

with corresponding $N^* = 6286$ and 6027 , respectively. Expecting g to be the typical survivor, we choose an initial population dominated by the *less fit* g' ,

so as to study the crossover behavior. From 10,000 runs with an initial fraction of only 1.67% of the *fitter* g , we compile histograms to provide $P(f, t)$, which is the probability (density) to find, at time t , the fitter genome (g) being a fraction f of the whole population. Though some runs end with g going extinct, it “wins” 97.5% of the time. To understand this crossover, it is not necessary to struggle with (6,7). An adequate description, at the “coarsest” level, comes from the most naive assumptions, i.e., postulating an equation for $f(t)$ with the appropriate fixed points ($f = 0, 1$): $df/dt = rf(1 - f)$, where r is a rate which can be fitted to the data. The solution to this equation is trivial and, with $r \simeq 0.0133$ in this case, the curve fits well the “ridge” in a contour plot of $P(f, t)$.

3.3 Competition between comparable genomes

For this study, we choose another pair of genomes:

$$g = 0000\ 0000\ 1100\ 0001\ 0101\ 0000\ 0000\ 0000 \quad (16)$$

$$g' = 0000\ 0001\ 1000\ 0000\ 1000\ 0000\ 1100\ 0000 \quad (17)$$

with effectively degenerate N^* 's, namely, 5344. Since mean-field theory predicts the presence of a fixed line, fluctuations should play a dominant role. The most naive guess would be that the system executes a “random walk” along this line. In Fig. 1, we present the results for $f(t)$ of several typical runs. In each graph, f is the fraction of the population made up of genome g , while the units of t are just “years.” All runs start with an equal mixture

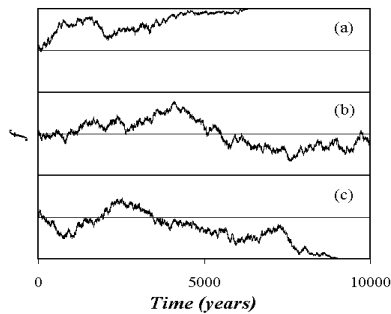


Fig. 1. Three typical runs in the competition of two similar genomes (without mutations), showing fractions as a function of time. In the upper/lower plot, g/g' “wins out.” In the middle plot, both are still present after 10K generations

of babies of the two genomes. At first glance, the time traces do resemble random walks. Of course, since the top(bottom) of the graphs represents the extinction of $g'(g)$, they might be random walks with absorbing boundaries (RWAB). Such walks are standard, described by $P_{RW}(f, t)$, which satisfy

$$\partial_t P_{RW} = D \partial_f^2 P_{RW} \quad (18)$$

with $P_{RW}(0, t) = P_{RW}(1, t) = 0$. Here, the diffusion constant, \mathcal{D} , can be related to the step size of the underlying walk. However, this picture proves to be too naive. Specifically, the late time behavior of the simulations turns out to be quite *distinct* from that of a simple RWAB.

Collecting data from 6200 runs of up to 85,000 years, we construct $P(f, t)$ from compiling histograms in f . As soon as one of the species goes extinct, the run is stopped. Thus, the single species components (i.e., $\delta(f)$ and $\delta(1 - f)$) are not represented in the figures below. Plots of $P(f, t)$ for an “intermediate time” ($t = 2K$) and a “late time” ($t = 20K$) are shown in Fig. 2a. Meanwhile, the standard solution for the RWAB, with $\delta(f - 0.5)$ as initial condition, is $P_{RW}(f, t) = 2 \sum_{\text{odd } n} (-1)^{(n-1)/2} \sin(n\pi f) e^{-\mathcal{D}(n\pi)^2 t}$. Thus, at late times, P_{RW} is proportional to $\sin(\pi f)$, as shown in Fig. 2b, in stark contrast to the *flat distributions* seen in Fig. 2a. Indeed, from $t \approx 5K$ on, P is essentially flat with a slowly decaying amplitude.

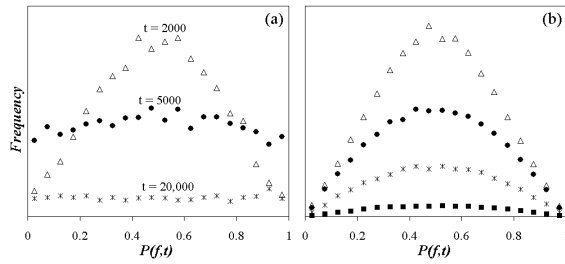


Fig. 2. (a) Histograms of 6200 runs are compiled for this $P(f, t)$ vs. fraction f , for various t 's. (b) Histograms of 50,000 runs for a RWAB

In an attempt to resolve this difference, the “walks” in the population dynamics are examined more closely. In particular, the step sizes, unlike those in RWAB, depend on f , decreasing towards the extremal points. With some thought, this phenomenon can be expected. Since the birth and death *rates* are the constants in our dynamics, smaller changes naturally accompany smaller populations. Known as Fisher Wright diffusion, random walks of this type were first introduced in the context of genetic evolution diffusion [2,12,14]. Instead of (18), the equation

$$\partial_t P_{FW} = \mathcal{D} \partial_f^2 \{f(1-f) P_{FW}\} \quad (19)$$

clearly incorporates variable step sizes. Many of the ingredients used by Fisher and Wright in their considerations are applicable to our model. Thus, it is not surprising that the data in Fig. 2a conform to this description. To be more specific, let us write the solution to (19) as

$$P_{FW}(f, t) = \sum_n A_n u_n(f) e^{-\lambda_n t} \quad (20)$$

where $u_n(f)$ and $-\lambda_n$ are the eigenfunctions and eigenvalues of the differential operator $\partial_f^2 f(1-f)$. It is easy to verify [13] that the u 's are just

Gegenbauer polynomials of order n , while $\lambda_n = \mathcal{D}(n+1)(n+2)$. Fixing the amplitudes A_n via $\delta(f - 1/2) = \sum_n A_n u_n(f)$ at $t = 0$, the full evolution is known. First, notice that the eigenfunction associated with the slowest decay is just $u_0(f) = 1$, which is in complete agreement with the flat distributions in Fig. 2a. We can further test the validity of (20) by projecting $P(f, t)$ onto the u_n 's and checking if the resultant $A_n(t)$'s decay exponentially. In Fig. 3, we show the behavior of the lowest non-zero pair: $A_{0,2}(t)$. It is clear that, not only are the exponential tails unambiguously present, the fitted decay rates are in excellent agreement with $\lambda_0:\lambda_2 = 1:6$.

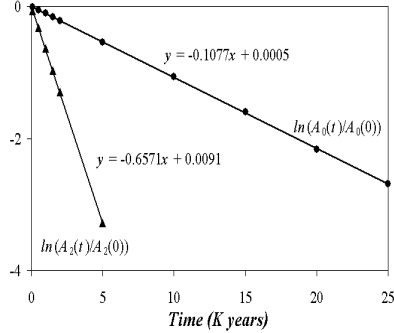


Fig. 3. Amplitudes $A_{0,2}(t)$, extracted according to the Fisher-Wright diffusion equation

Beyond this “coarsest” phenomenological approach, we turn to a slightly more detailed description, from which we can *derive* the Fisher-Wright equation for $P(f, t)$. Focusing on a population with two genomes, but *without age structure*, we consider the numbers of each species: (N, M) . Similar to the considerations for RWAB, we will use continuous time for simplicity and propose a master equation for $P(N, M; t)$. There are two types of terms here: death rates proportional to the total population $(N + M)$, and constant birth rates. Absorbing an overall rate into the scale of t and fixing the only parameter to reproduce a steady state total population at \mathcal{N}^* , we arrive at the unique equation:

$$\begin{aligned} \partial_t P(N, M) = & (N + M + 1) [(N + 1)P(N + 1, M) + (M + 1)P(N, M + 1)] \\ & + \mathcal{N}^* [(N - 1)P(N - 1, M) + (M - 1)P(N, M - 1)] \\ & - (N + M)(N + M + \mathcal{N}^*) P(N, M) . \end{aligned} \quad (21)$$

It is straightforward to derive evolution equations for the averages $\langle N \rangle_t \equiv \sum_{N, M} N P(N, M; t)$ and $\langle M \rangle_t$, though these will involve, as usual, higher moments. However, taking the mean-field approximation, these reduce to equations for $\langle N \rangle_t$ and $\langle M \rangle_t$, which are conceptually clear. If $\mathcal{N}^* \gg 1$, we may divide (21) by it and take the continuum limit by defining

$$s \equiv (N + M) / \mathcal{N}^* \quad \text{and} \quad f \equiv N / (N + M) .$$

Deferring all mathematical details to another publication [15], we simply report that the resultant is a second order, *separable* partial differential equation (for $P(s, f; t)$), in which f enters only via the operator: $\partial_f^2 f (1 - f)$. This analysis shows that an equation like (19) for $P(f, t) \equiv \int ds P(s, f; t)$ is a natural consequence of the slightly less “coarse” approach in (21). This latter description is similar to the “bivariate model” of Feller [12].

Finally, for completeness, we provide the most “microscopic” description of the entire process: a master equation for $P(\{n_g(\alpha)\}; t)$ in discrete time. Exploiting the simplified notation (11,12), we have

$$\begin{aligned}
 P(\{n_\alpha, m_\beta\}, t + 1) &= \sum_{\{k_\alpha, \ell_\beta\}, n_{D+1}, m_{D'+1}} P(\{k_\alpha, \ell_\beta\}, t) \\
 &\times \prod_{\alpha, \beta} \left[\binom{k_{\alpha-1}}{n_\alpha} V^{n_\alpha} (1 - V)^{k_{\alpha-1} - n_\alpha} \binom{\ell_\beta}{m_\beta} V^{m_\beta} (1 - V)^{\ell_\beta - m_\beta} \right] \\
 &\times \prod_{\alpha, \beta} \left[\binom{n_\alpha}{\tilde{n}_\alpha} w_\alpha^{\tilde{n}_\alpha} (1 - w_\alpha)^{n_\alpha - \tilde{n}_\alpha} \binom{m_\beta}{\tilde{m}_\beta} u_\beta^{\tilde{m}_\beta} (1 - u_\beta)^{m_\beta - \tilde{m}_\beta} \right] \\
 &\times \sum_{\{\tilde{n}_\alpha, \tilde{m}_\beta, B_\alpha, C_\beta\}} \delta \left(n_0 - \sum_\alpha \tilde{n}_\alpha F_\alpha \right) \delta \left(m_0 - \sum_\beta \tilde{m}_\beta G_\beta \right). \quad (22)
 \end{aligned}$$

Here V actually stands for $\exp[-\sum_\alpha (k_\alpha + \ell_\alpha)/\mathcal{N}_0]$, to conform with our simulation studies. The products range over the ages from 1 to their respective deaths: $D \equiv D_g$ and $D' \equiv D_{g'}$. Note that the presence of n_{D+1} and $m_{D'+1}$ on the right does *not* imply that we have survivors beyond the death ages. These variables are introduced in order to account properly for individuals of the maximum age surviving the competition to reproduce, *before* dying. At the risk of being too pedantic, let us elucidate each line: the second shows the probability of n_α, m_β out of $k_{\alpha-1}, \ell_{\beta-1}$ adults surviving the year; the third represents the probability of $\tilde{n}_\alpha, \tilde{m}_\beta$ out of n_α, m_β adults becoming parents; and the last selects only those combinations of babies that match n_0, m_0 . Note that the probabilities $w_g(\alpha), w_{g'}(\beta)$ and fecundities $F_g(\alpha), F_{g'}(\beta)$ from Eqn (1) have been abridged to w_α, u_β and F_α, G_β , respectively.

This equation generalizes the one proposed earlier[9] for populations with a single species. Clearly, this will lead to much more detailed statistical information than either $P(f, t)$ or $P(N, M; t)$, such as (oscillatory) correlations between various ages. Needless to say, we can expect anti-correlations between the species (in an overall sense). However, the full description can provide more interesting aspects, especially if one species is much more highly productive but much shorter lived. Beyond the scope of this paper, the results of these studies will be published elsewhere[7,15]. Instead, we turn to the effects of mutation on a population with two species.

4 Two species competition, with mutations

In this section, we study the dynamics of a population with two genomes when mutations are present. In typical studies of the Penna model [4,6,5] where mutations consist of changing arbitrary bits (at some fixed rate), all 2^L genomes can exist so that only some crude properties of how they compete are feasible (unless relatively short strings are used). With the eventual goal of distributions in the full g -space in mind, we restrict mutations in this study to be only *between* the two species. Though somewhat artificial, we believe this exercise is a worthwhile detour, for providing a clear and simple picture of the effects of mutations. For this study, μ obviously represents a parameter different from that in the Penna model. Here, $(1 - \mu)$ is simply the fraction of babies born with the genome of the parent, while the rest inherit the genome of the competitor.

Clearly, if $\mu \neq 0$, neither species will go extinct and the fixed point structure must be different from that described above. It is interesting to examine the combined effects of competition and mutations, over a range of μ , on genomes of similar “fitness” (N^*) and very different ones. In particular, we can expect, in simulations, non-trivial steady state distributions, i.e., $P^*(f) \neq c\delta(f) + (1 - c)\delta(1 - f)$.

4.1 The mean-field description

First, we begin with the deterministic mean-field approach. Instead of (6,7), the evolution equations now read:

$$\begin{aligned} n_0(t+1) &= V(\mathcal{N}) \sum [(1 - \mu) p_\alpha n_\alpha(t) + \mu q_\alpha m_\alpha(t)] \\ m_0(t+1) &= V(\mathcal{N}) \sum [\mu p_\alpha n_\alpha(t) + (1 - \mu) q_\alpha m_\alpha(t)] \\ n_{\alpha+1}(t+1) &= V(\mathcal{N}) n_\alpha(t) \quad \text{and similar for } m_\alpha. \end{aligned} \quad (23)$$

The analysis leading to the fixed point is similar to the no-mutation case. Provided the productivities are neither too low nor too high, there is a *unique* solution, within the range $v \in (0, 1)$, to

$$\frac{1}{p(v)} + \frac{1}{q(v)} = 2 - \frac{1}{\mu} \quad (24)$$

where $p(v) \equiv 1 - \sum p_\alpha v^{\alpha+1}$ and $q(v) \equiv 1 - \sum q_\alpha v^{\alpha+1}$. To appreciate the qualitative features of this solution, suppose $v_g < v_{g'}$, where $p(v_g) = 0 = q(v_{g'})$. Then $v_g \leq v \leq v_{g'}$, i.e., it lies *between* the values associated with each of the two species. At this point, $p < 0$ and $q > 0$. Note that, with μ being strictly zero, both $p(v) = 0$ and $q(v) = 0$ are allowed, these being the conditions for the single genome fixed points. However, *in the limit* $\mu \rightarrow 0$, only the “fitter” species (smaller v , larger N^*) is picked out. Another interesting point is $\mu = 1/2$, corresponding to babies having no “memory” of

their parents' genetics! In this case, the intuitively reasonable result emerges after a little algebra, i.e., v satisfies $\sum \bar{p}_\alpha v^{\alpha+1} = 1$, where \bar{p}_α is just the average productivity: $(p_\alpha + q_\alpha)/2$.

Once v is known, we have the total population N^* and the age distributions $n_\alpha^* = v^\alpha n_0^*$; $m_\alpha^* = v^\alpha m_0^*$. The relative populations can be determined from adding the first two equations in (23):

$$\frac{n_0^*}{m_0^*} = -\frac{q(v)}{p(v)}, \quad (25)$$

which is positive. Note that, unless the two genomes have the same longevity, the ratio of the totals of each species will *not* be the same as (25). Instead, we have

$$\frac{N^*}{M^*} = \frac{-q(v) (1 - v^{1+D})}{p(v) (1 - v^{1+D'})}. \quad (26)$$

4.2 Simulation results

Turning to simulations, we begin with the simplest case: $\mu = 1/2$, which is an analog of a thermal system at infinite temperature. Since all newborns are randomly assigned one of the two genomes, a *naive* expectation is that the stationary distribution of the fraction f to be a Gaussian centered at $f = 1/2$. However, in simulations using the pair with different N^* 's: (14,15), we found the distribution to be centered slightly towards the one with the *smaller* N^* (or less "fit")! The explanation lies in the difference in the longevity: g' survives six years longer in this case. Indeed, the mean-field description (26) already shows the dependence of population ratios on their death ages. As μ increases from zero to $1/2$, longevity plays a more important role than productivity in determining "fitness." ¹ To see this effect in a sharper contrast, we perform simulations with g' above (15) competing against an "alien" species (i.e., from *outside* the Penna-Desai paradigm). Specifically, the latter is short-lived ($D = 2$) but precocious (large and early reproductivities: $p_1 = 5$, $p_2 = 3$, so that $N_{alien}^* \sim 3N_{g'}^*$). Fig. 4a shows the steady state $P^*(f)$ for several values of $\mu \in [0, 1/2]$. The effect of longevity is dramatically displayed, as we observe the peaks of the distributions move with the mutation rate, from ~25% to 100% "aliens". Finally, notice that all these distributions are single peaked, so that the "transition" from a productivity dominated population to a longevity dominated one is most likely a simple cross-over (or, at the most, a continuous transition).

¹ * For a recent article on the issue of fitness versus longevity, see W. Hwang, P. L. Krapivsky, and S. Redner, adap-org/9912004. Being a more complex model, it displays a more subtle phenomenon than the one described here.

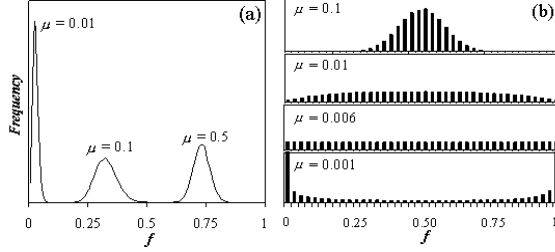


Fig. 4. (a) Steady state distributions when g' competes against a short-lived productive species (f being fraction of the latter). (b) Steady state distributions when g' competes against g

Next, we turn our attention to competition between species with comparable N^* 's. Specifically, we chose

$$g = 0000\ 0000\ 1000\ 0000\ 0001\ 1010\ 0010\ 0000 \quad (27)$$

$$g' = 0000\ 0000\ 1000\ 0000\ 0100\ 0000\ 1110\ 0000. \quad (28)$$

Here, mean-field theory predicts the presence of a fixed line rather than a fixed point, so that fluctuations will be very important. Since we expect to need exceedingly long runs, we lower \mathcal{N}_0 to 2000, so that both have $N^* \cong 569$. Here, we choose to focus on two phenomena: (a) the steady state distributions $P^*(f)$ (as μ is lowered from $1/2$ to 0) and (b) the lifetimes of single species dominated states (for very small μ 's).

For part (a), we used a single run for each μ , starting with equal numbers of the two species and lasting 10^7 years. To avoid long transients in age-structures, we endow each species with its steady state distribution of ages (i.e., (10)) initially. In case there are other transients, we discard the first 100 years before collecting data. Also, to minimize correlations while compiling histograms for $P^*(f, \mu)$, we sampled the system at random intervals between 5 and 15 years. Clearly, we can expect $P^*(f)$ to be symmetric about $1/2$, and that, for systems with symmetric initial conditions, it remains symmetric throughout the run. Unlike the case for very different species, $P^*(f)$ for $\mu = 1/2$ is *trivially* Gaussian. On the other hand, we have seen that, for $\mu = 0$, the (symmetric) steady state distribution is just the sum of two δ -functions: $[\delta(f) + \delta(1-f)]/2$. Thus, as the rate is lowered, $P^*(f)$ must turn from a single peak to a bi-modal distribution. In Fig. 4b, we show $P^*(f)$ for four rates $\mu = 0.1, 0.01, 0.006$, and 0.001 . The first shows the presence of a well-shaped Gaussian. The second illustrates that $P^*(f)$ is single peaked down to the point where the “finite size” of the range of f comes into play. In the third plot, we find a μ which leads to an essentially *flat* distribution. Finally, the last shows the expected bi-modal state.

“Transitions” from single- to double-peaked distributions are reminiscent of $P_I(m)$, the distribution for the magnetization, m , in an equilibrium Ising model (with nearest neighbor interaction J). As T is lowered from ∞ to 0, it also begins as a trivial Gaussian (binomial) centered on $m = 0$ and ends as $[\delta(1+m) + \delta(1-m)]/2$. The notion that μ plays a role similar to

that of T is entirely plausible, since both are responsible for “randomizing” the system. Needless to say, the analogy ends quickly. First, our $P^*(f)$ is certainly inconsistent with the Ising model in $d \geq 2$, for which a second order phase transition occurs at a non-zero critical temperature, so that $P_I(m)$ is bi-modal, with peaks *away from* the end points $m = \pm 1$, for most of $T < T_c$. Though resembling more the Ising model in *one* dimension and of *finite length* L , our $P^*(f)$ differs in significant details. It turns out that, as T is lowered beyond $O(J/k_B \ln L)$, $P_I(m)$ becomes *tri*-modal, with peaks at $m = 0, \pm 1$. At no point is $P_I(m)$ flat or bi-modal!

While the role of μ resembles that of T (or $e^{-2J/k_B T}$) in the $d = 1$ Ising model, the parameter comparable to L is \mathcal{N}_0 . By carrying out simulations with $\mathcal{N}_0 = 20$ and 2000, we find that the distributions become nearly flat at $\mu \simeq 0.06$ and 0.0007, respectively. Based on this preliminary finding, we conjecture that $\mu \mathcal{N}_0$ is the scaling variable, similar to $e^{-2J/k_B T} L$ for $P_I(m)$. An extensive Monte Carlo study to confirm (or to disprove) this hypothesis is in progress and will be reported elsewhere [15].

Lastly, we turn to another phenomenon, which also has an analogue in the finite Ising model in $d = 1$. For extremely low temperatures (i.e., $T \ll J/k_B \ln L$), the Ising system is confined mainly to one of the two extreme states ($m = \pm 1$). Nevertheless, for any positive T , the system will make excursions and flips into the other state. Similarly, we can expect our model, with minute amounts of μ , to make periodic transitions from a population dominated by one genome to that dominated by the other. One measure of this phenomenon is the average time between the transitions: τ . Since the transitions are relatively “sharp” (in time), we choose to define the *completion* of a transition when the originally dominating species dwindles to 2% of the total population. Expecting to need exceedingly long runs, we considered only three values of \mathcal{N}_0 : 200, 2000, and 20,000. Fixing a μ , we start the system with $N_g, N_{g'} = 10, 600$ and let it evolve until 300 transitions occur. Not surprisingly, the runs last much longer for smaller μ . Over the chosen range of μ (from 0.03 to 3×10^{-6}), the runs range from about 2×10^5 to

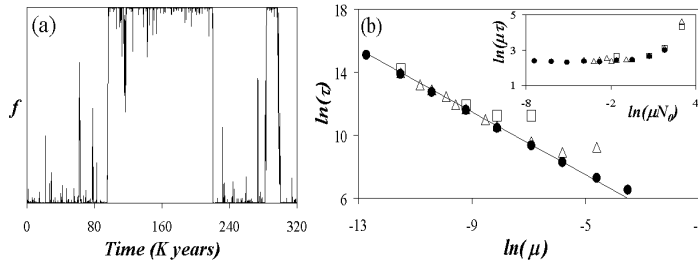


Fig. 5. (a) A typical run with $\mu = 10^{-4}$, showing flipping between g - and g' -dominated populations. (b) Log-log plot of average lifetime vs. μ for three \mathcal{N}_0 's. A line of slope -1 is drawn to guide the eye. Inset: Scaling plot of $\mu\tau$ vs. $\ln(\mu\mathcal{N}_0)$

over 10^9 years! To illustrate our data, we show a small section of a typical run in Fig. 5a, displaying f over 3.2×10^5 years, for $\mu = 10^{-4}$. (Contrast the time scale in this figure with that in Fig. 1!) Our expectation is realized: f tends to stay near the extremes until a large fluctuation takes it to the region near $1/2$. Then, competitions similar to the $\mu = 0$ case return the system to the extremal points. For each run, 300 lifetimes are recorded and the average τ computed. From a plot of $\ln(\tau)$ vs. $\ln(\mu)$ for the three \mathcal{N}_0 's (Fig. 5b), we find that the curves merge roughly into a single straight line of slope -1 , once μ drops below a value of $O(1/\mathcal{N}_0)$. A crude test of the scaling hypothesis is shown in the inset of Fig. 5b, where $\mu\tau$ is plotted versus $\ln(\mu\mathcal{N}_0)$. The quality of data collapse suggests that $\mu\tau$ is a function of $\mu\mathcal{N}_0$ only. Obviously, a more extensive Monte Carlo study will clarify this issue also. In the next subsection, we will show how the data are consistent with the coarse-grained approach.

4.3 Analytic approaches beyond mean-field

On the analytic front, it is simple to generalize the above three levels of description beyond mean-field theory, from least refined $P(f, t)$, to the most detailed $P(n_\alpha, m_\beta; t)$. As in the previous section, to capture the essentials of competition, it suffices to postulate an appropriate master equation for $P(f, t)$. Following Fisher and Wright [2,12], the generalization of (18) is

$$\partial_t P_{FW} = \{\partial_f \mathcal{B} (2f - 1) + \mathcal{D} \partial_f^2 f (1 - f)\} P_{FW} \quad (29)$$

where the extra bias (\mathcal{B}) is to describe the effects of μ . This equation also appeared in many other contexts [13]. For example, rewritten in terms of $x \equiv 2f - 1$, it is the Fokker-Planck equation for a simple harmonic oscillator (with restoring force $-(\mathcal{B} - \mathcal{D})x$) subjected to a special kind of multiplicative noise. As in the case of no mutations, (29) is also completely soluble [13]. Thanks to the presence of \mathcal{B} , there is now a non-trivial steady state: $P_{FW}^* \propto [f(1-f)]^{\mathcal{B}/2\mathcal{D}-1}$. The most important aspect is that P_{FW}^* displays a “transition,” as \mathcal{B} is lowered, from being single peaked to bi-modal. Furthermore, at the “critical” value $\mathcal{B}_c = 2\mathcal{D}$, the distribution is *flat*. Thus, this approach should be successful in describing our simulations.

Instead of fitting another parameter, let us turn to the “intermediate” level description, based on $P(N, M; t)$. Modifying Eqn (21) to include mutations is straightforward. Deferring the mathematical details once again [15], here we only report that the continuum limit of this generalization contains the differential operator

$$\partial_f^2 [f(1-f)] + \mu \mathcal{N}^* \partial_f (2f - 1) \quad (30)$$

for f . In other words, this approach allows us to identify the ratio

$$\mathcal{B}/\mathcal{D} = \mu \mathcal{N}^* . \quad (31)$$

Since $\mathcal{N}^* \propto \mathcal{N}_0$, we see that this description justifies the scaling variable $\mu\mathcal{N}_0$. However, the precise value for the critical μ from simulations is about 50% higher than $2/\mathcal{N}^*$. The discrepancy may be a signal that this “mesoscopic” approach is indeed too crude. Perhaps it is similar to the Hubbard-Stratonovich[16] method for obtaining a Landau-Ginzburg Hamiltonian for Ising models, in which the precise critical temperatures are missed. Of course, we can go to the fully “microscopic” description, writing a generalization of (22) for $P(\{n_\alpha, m_\beta\}, t)$. In principle, this approach should predict all details. However, since we have not yet succeeded in extracting useful analytic information, we will spare the reader of this complicated equation.

Proceeding to the problem of predicting τ , the average time of single species dominance, we note that it is far from being solved, even though an exact solution for $P^*(f)$ is available. Indeed, finding τ is a problem of first passage times, which requires much more analysis, as is well known in simple random walks [17], than finding stationary distributions. Work is in progress to study the Fisher-Wright process with discrete space-time steps. Hopefully, the first passage time problem is soluble within such a framework, so that we have some quantitative understanding of the scaling displayed in Fig 9. Perhaps there is some relationship to the problem of persistence [18], where non-trivial properties are unfathomed from linear stochastic processes with uncorrelated white noise!

5 Summary and outlook

We have studied, with both Monte Carlo simulations and analytic methods, the effects of simple competition between just two species with age structure. For definiteness, we used the recently introduced Penna-Desai model [4,6], in which individuals age, survive, reproduce and die according to specific rules. In the absence of mutations, one species eventually dominates the entire population. For pairs with different “fitness,” the evolution is well described by a mean-field approach, based in the most naive assumptions. For comparable species, however, mean-field theory fails, so that we must consider probabilistic approaches. At the coarsest level, we focus on only $P(f, t)$, the probability that we find a fraction f of our population being one species, at time t . Without mutations, the steady state is a trivial linear combination of the single species states: $\delta(f)$ and $\delta(1 - f)$. However, non-trivial late time properties were observed. Specifically, instead of being a simple random walk (on $f \in [0, 1]$) with absorbing boundary conditions, our $P(f, t)$ fits well into Fisher-Wright diffusion [2,12]. This behavior can be derived from an intermediate level description, which relies on a reasonable master equation for $P(N, M; t)$, the probability that our population is composed of N individuals of one species and M of the other, at time t . Finally, we provide the master equation for $P(n_\alpha, m_\beta; t)$, a complete “microscopic” description for $n_\alpha(m_\beta)$ individuals of age $\alpha(\beta)$. Next, we turn to the effects of mutation, with which

a fraction μ of newborns bear the genome of the other species. Now, a non-trivial steady state $P^*(f)$ exists. Although the functional form appears to fit well into the Fisher-Wright scheme, the dependence of the phenomenological parameters on μ and N_0 cannot be predicted, either at the coarsest or the intermediate level. We believe that this dependence must arise from the full “microscopic” description, though such a derivation remains to be carried out. In some ways, the qualitative features of two-species competition are comparable to the Ising model in $d = 1$, with μ playing the role of temperature and $g - g'$ the role of the magnetic field. It would be interesting to see if this analogy can help in the understanding of the much more complex problem of multi-species competition.

Going beyond this simple model of only two species, we may consider many interesting generalizations. As they take into account more realistic aspects of real populations, they become more complex and intractable. Here we only mention a few.

The most obvious extension of our study is the full Penna’s model, which allows specific mutations between the 2^L species. Even staying within a mean-field description, the steady state is non-trivial. In particular, we are not aware of any analytic solutions to the general case, which can provide some insight into the general properties. Each case (i.e., specific L and rule of mutation) must be solved separately. From a simple $L = 3$ model, we find that the connectivity between g ’s (due to mutations) plays a very important role in determining the most probable steady state distribution in N_g , to the extent that these are completely unrelated to the N_g^* ’s. Clearly, much more work is needed to elucidate these issues.

So far, we (and many others working with the Penna model) have incorporated competition in the most simplistic manner, namely, through a Verhulst factor depending only on the total number of individuals \mathcal{N} . In general, differentiated competition can be introduced through distinct factors for each species, with dependence on the numbers of all other species: $V_g(\{N_{g'}\})$. Even more realistic is to include dependence on the age structures: $V_g(\{n_{g'}(\alpha)\})$, so as to account for, e.g., competition for food being stronger from N adults than from the same number of babies. Similarly, we have studied only a model where the birth rates are environment independent. Generalizing to $p_g(\alpha; \{n_{g'}(\alpha')\})$ would include all predator-prey models. Clearly, “internal” rates of declining survivability can be included (e.g., [19]) also. Branching out in another direction, we could study the limit of continuum α . Casting such models in the field theoretic language, we would be faced with long-range “interactions” (between a range of α and the end point $\alpha = 0$) with multiplicative noise [9]. Thus, despite the fact that this is “just a $d = 1$ field theory”, there is the tantalizing possibility that a real phase transition can exist. Finally, we note that many further generalizations could be included, such as sexual reproduction and spatial-temporal dependence (e.g., diffusion, migration, advection). Though simulations are easy to perform, theoretical

understanding is, understandably, much more difficult. Ours is but a minute step towards the ultimate goal of gaining some insight into evolution in real ecosystems [20].

Acknowledgements

We thank R. Desai for introducing us to Penna's model and B. Schmittmann, A.J. McKane, M. Howard for illuminating discussions. We are indebted to H. Spohn for pointing out the possible connection between the stochastic process here and Fisher-Wright diffusion and to C. Doering for informing us of ref. [13]. One of us (RKPZ) is grateful to H.W. Diehl for his hospitality during a stay at the University of Essen, which substantially facilitated this research. This work is supported in part by a grant (DMR-9727574) from the National Science Foundation.

References

1. P. F. Verhulst, *Corres. Math. et Physique* **10**, 113 (1838).
2. R.A. Fisher, *Proc. R. Soc. Edinb.* **42**, 321 (1922) and S. Wright, *Genetics* **16**, 97 (1931). For a more recent treatment, see, e.g., [12].
3. A.J. Lotka, *Elements of Mathematical Biology*, (Dover, NY, 1956); V. Volterra, in R.N. Chapman, *Animal ecology*, (McGraw Hill, NY, 1931).
4. T.J.P. Penna, *J. Stat. Phys.* **78**, 1629 (1995).
5. S. Moss de Oliveira, D. Alves, J.S. Sá Martins, cond-matt/0004072; H. Puhl, D. Stauffer, S. Roux, *Physica* **A221**, 445 (1995); T.J.P. Penna, S. Moss de Oliveira, D. Stauffer, *Phys. Rev.* **E52**, R3309 (1995).
6. R.C. Desai, F. James, E. Lui, *Theo. Biosci.* **118**, 97 (1999).
7. R.J. Aсталos, Dynamics of Competition using a Bit String Model with Age Structure and Mutations, Ph.D. thesis, Virginia Tech, 2001.
8. J.D. Murray, *Mathematical Biology*, (Springer, Berlin, 1993).
9. M. Howard and R.K.P. Zia, *Int. J. Mod. Phys.* **B** (2001); cond-mat/0001172.
10. R.M. May, *Science* **186**, 645 (1974).
11. H.M. Taylor and S. Karlin, *An Introduction to Stochastic Modeling*, 3rd Ed., p.464 (Academic Press, San Diego, 1998).
12. W. Feller, *Proc. Second Berkeley Symp. on Math. Stat. and Prob.* p.227 (1951).
13. C. Doering, *Phys. Lett.* **A122**, 133 (1987).
14. W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. 2, (Wiley, NY, 1966).
15. R.J. Aсталos and R.K.P. Zia, to be published.
16. J. Hubbard, *Phys. Rev. Lett.* **3**, 77 (1958) and R.L. Stratonovich, *Doklady* **115**, 1097 (1957) [English translation: *Soviet Phys-Doklady*, **2**, 416 (1958)].
17. B. D. Hughes, *Random Walks and Random Environments*, (Oxford, Oxford, 1995).
18. S.N. Majumdar and A.J. Bray, cond-mat/0009439 (2000).
19. Z.F. Huang and D. Stauffer, cond-mat/0007473 (2000).
20. B. Drossel, cond-mat/0101409, to be published in *Advances in Physics*. See also, e.g., B. Drossel, P.G. Higgs, and A.J. McKane, cond-mat/0002032, to be published in *J. Theo. Biol.*