Recombination and Bistability in Finite Populations

Lionel Barnett

Centre for the Study of Evolution
Centre for Computational Neuroscience and Robotics
School of Cognitive and Computing Sciences
University of Sussex
Brighton BN1 9QH, UK

lionelb@cogs.susx.ac.uk

September 29, 1998

Abstract

In this paper we analyse the phenomenon of "bistability" in finite population evolutionary dynamics, especially with regard to recombination. Bistability, where the steady-state population distribution depends on the initial state of the population, has recently been observed in an (infinite population) quasi-species model of viral recombination (Boerlijst et al., 1996). We analyse a comparable finite population model using a birth and death process due to (Moran, 1958). Bistability (or its stochastic analogue) is revealed in the bimodality of the stationary probability distribution of the birth and death process and long mean transition times between the modes. These effects are demonstrated to be exaggerated by recombination.

1 Introduction

In (Boerlijst et al., 1996) a mathematical model for an asexual (haploid) quasi-species (Eigen et al., 1989) evolving with recombination is introduced to study recombination in retro-virus populations. The model is analyzed on several simple fitness landscapes. A striking feature of the model is the appearance of "bistability" or "hysteresis" in the steady-state population distribution for particular combinations of mutation rate and recombination rate; i.e. the steady-state distribution of genotypes depends on the distribution of genotypes in the initial population. This phenomenon has also been observed in various diploid models, both with and without recombination. In (Boerlijst et al., 1996) bistability in their (infinite-population, hence deterministic) model is explained in terms of bifurcation of the differential equations describing the time-evolution of the quasi-species. A recent empirical study (Ochoa and Harvey, 1999) suggests strongly that many qualitative features of the infinite-population model are preserved in the corresponding finite-population dynamics. In this paper we investigate bistability in finite-population stochastic population dynamics. The model we use is based on a birth and death process originally devised by (Moran, 1958) and previously deployed in a situation analogous to ours (but without recombination) by (Nowak and Schuster, 1989).

2 The Moran Model

In (Moran, 1958) a model for the evolution of a fixed-size finite population of genotypes was introduced, based on the idea of fitness as the expected (reproductive) lifetime of a genotype. Here we extend the model to arbitrary fitness landscapes and to include recombination.

Let $\boldsymbol{Q}^\nu$ represent the ν-dimensional binary hypercube; i.e. an element of $\boldsymbol{Q}^\nu$ is a binary sequence of length ν, which we identify with a haploid genotype. We specify a fitness landscape on $\boldsymbol{Q}^\nu$ by assigning to each $g \in \boldsymbol{Q}^\nu$ a (real-valued) *fitness* $f(g) > 0$. Consider a population comprising N such genotypes. We may identify such a population with an integer vector $\boldsymbol{n} = (n_g \mid g \in \boldsymbol{Q}^\nu)$, where n_g represents the number of copies of genotype g in the population, $n_g \geq 0 \ \forall g$ and $\sum_{g \in \boldsymbol{Q}^\nu} n_g = N$. We now define a *birth-death event* on the population $\boldsymbol{n}$ to be a transformation of $\boldsymbol{n}$ into a new population $\boldsymbol{n}'$ as follows: a copy of some genotype g_1 "dies" and a copy of another (possibly the same) genotype g_2 is "born". In terms of the population vectors we have $n'_{g_1} = n_{g_1} - 1$ and $n'_{g_2} = n_{g_2} + 1$; the population size thus remains constant.

Suppose now that we have a stochastic process $\{\boldsymbol{n}(t) \mid t \geq 0\}$ of populations (t represents a continuous time parameter) that evolves according to the following scheme: in any time interval $[t, t+h]$ the probability that a copy of genotype g dies is given by (Moran, 1958):

$$\mathbf{P}\left(\textit{a copy of } g \textit{ dies in the interval } [t, t+h]\right) = \frac{\theta}{f(g)} \frac{n_g(t)}{N} h + o(h) \tag{1}$$

where θ is a fixed timescale parameter. It is straightforward to verify that the "lifetime" of a genotype g is exponentially distributed with expectation $\frac{N}{\theta} f(g)$. A death triggers an immediate birth, thus defining a birth-death event. Candidates for a birth are selected as follows: with probability $1 - \rho$ the birth is *asexual* and with probability ρ *sexual*, where $0 \leq \rho \leq 1$ is the recombination rate. For asexual reproduction a parent is selected uniformly at random and with replacement from the population. The offspring is taken to be a copy of the parent mutated with per-allele probability μ where $0 \leq \mu \leq \frac{1}{2}$ is the *(per-allele) mutation rate*. In the sexual case two parents are independently selected uniformly at random and with replacement from the population. The parents are mated by *uniform recombination*; i.e. independently for each locus on the genotype, one of the parents is chosen at random and its allele (0 or 1) becomes the allele of the offspring at that locus. After recombination the offspring is mutated with mutation rate μ as in the asexual case. It may also be verified that the expected number of offspring of a given genotype during its lifetime is proportional to its fitness. To see this, note that for a given genotype the times between its successive selections as a parent are (identically and independently) exponentially distributed. Thus the number of offspring of a genotype from its birth up to a given time t constitutes a Poisson process (Stirzaker, 1994). From this and the exponential distribution of lifetimes the result follows by a straightforward calculation. In this sense, selection in the Moran model is *fitness proportional*.

It is clear that $\{\boldsymbol{n}(t) \mid t \geq 0\}$ thus defined is a Markovian birth and death process with state space the (vast!) set of all possible populations of size N. Because of the huge size and awkward structure of the state space it is difficult to say anything useful about such a process. In the next section we specialise to a specific simple fitness landscape and, with the help of some judicious approximations, reduce the state space to a tractable form.

3 The Single-peak Fitness Landscape

We specify a single-peak fitness landscape as follows: all genotypes have fitness 1 except for a single genotype, the "peak" or *optimal genotype*[1]. which has fitness σ where $\sigma > 1$ is the *selection coefficient*. Without loss of generality we take the optimal genotype to be the sequence of ν zeroes. For $\alpha = 0, 1, 2, \ldots, \nu$ let us define (Eigen et al., 1989) the *error class* $E_\alpha \subset \boldsymbol{Q}^\nu$ to be the set of all genotypes Hamming distance α from the optimum; i.e. with exactly α bits set. The E_α with $\alpha > 0$ are said to constitute the *error tail*.

Given a Moran birth and death process as described above on such a landscape, we will be interested in the number $X(t)$ of copies of the optimal genotype present in the population at time t. It would be convenient if the process $\{X(t) \mid t \geq 0\}$ were also a Markov process - unfortunately it is clear that the Markov property does not hold. This is because the probability that a non-optimal genotype (i.e. a genotype in the error tail) mutates to the optimal genotype depends on the distribution of genotypes over the error classes; without knowing this distribution we cannot know the probability that a birth will be optimal. In (Nowak and Schuster, 1989) this issue is addressed by making a "maximum entropy" approximation; specifically, it is assumed that a genotype in the error tail is as likely to be any one (non-optimal) genotype as another; i.e. that the distribution of genotypes in the error tail is always uniform random. This implies that at any time, $\alpha = 0, 1, 2, \ldots, \nu$:

$$\mathbf{P}\left(g \in E_\alpha \mid g \in \textit{error tail}\right) = \kappa \binom{\nu}{\alpha} \tag{2}$$

and

$$\mathbf{P}\left(\textit{an arbitrary bit of } g \textit{ is set} \mid g \in E_\alpha\right) = \frac{\alpha}{\nu} \tag{3}$$

where, following (Nowak and Schuster, 1989), we have set $\kappa \equiv \dfrac{1}{2^\nu - 1}$. Note that (3) implies the abscence of *linkage disequilibrium* between loci.

Under the assumptions (2) and (3) $\{X(t) \mid t \geq 0\}$ is indeed a Markovian continuous-time birth and death process with state space the set of integers from 0 to N and retaining barriers at 0 and N. If the mutation rate μ is non-zero then the process is also *irreducible* (Stirzaker, 1994) and thus has a unique stationary distribution (Karlin and Taylor, 1975). Such processes are quite well-understood and tractable to analysis; the question remains as to how well our approximation agrees with the original Moran birth and death process. It is, in fact, well-known that (2) does not hold in general (Nowak and Schuster, 1989; Boerlijst et al., 1996; Ochoa and Harvey, 1999). In particular, at low mutation rates the distribution of genotypes over the error tail is skewed towards the optimum - this is more or less the defining characteristic of a quasi-species! Furthermore, (3) will not in general hold due to *neutral drift* of the population (Kimura, 1983; Crow and Kimura, 1970; Derrida and Peliti, 1991) *within* the individual error classes. These issues will be addressed in a future paper. Suffice at this stage to note that preliminary research suggests that the behaviour of the model using the maximum entropy approximation agrees surprisingly well with the full model over a wide range of parameter values and that, in particular, it appears to preserve at least qualitatively the features addressed in this paper [2].

[1] Generally known as the *master sequence* in the quasi-species literature.

[2] It is also worth pointing out that near the *error threshold* (Eigen et al., 1989) the approximation (2) becomes more accurate. This is reflected in the accuracy of the error threshold approximation calculated in (Nowak and Schuster, 1989).

4 Analysis of the Birth and Death Process

We are now in a position to calculate the infinitesimal generators (Karlin and Taylor, 1975) of the simplified birth and death process $\{X(t) \mid t \geq 0\}$. To this end it suffices to know the probabilities:

$$\begin{aligned} m_1 &= \mathbf{P}\,(\textit{optimum mutates to optimum}) \\ m_2 &= \mathbf{P}\,(\textit{non}-\textit{optimum mutates to optimum}) \\ r_{11} &= \mathbf{P}\,(\textit{optimum recombined with optimum is optimum}) \\ r_{12} &= \mathbf{P}\,(\textit{optimum recombined with non}-\textit{optimum is optimum}) \\ r_{22} &= \mathbf{P}\,(\textit{non}-\textit{optimum recombined with non}-\textit{optimum is optimum}) \end{aligned} \tag{4}$$

Using (2), (3) and the definition of uniform recombination a straightforward if tedious computation yields:

$$\begin{aligned} m_1 &= Q \\ m_2 &= \kappa(1-Q) \\ r_{11} &= 1 \\ r_{12} &= \eta \\ r_{22} &= \kappa(1-2\eta) \end{aligned} \tag{5}$$

where we have set $Q \equiv (1-\mu)^\nu$ and $\eta \equiv \dfrac{(\frac{3}{2})^\nu - 1}{2^\nu - 1}$. The infinitesimal generators $\lambda_i,\ \mu_i$ of the birth and death process are defined by (Karlin and Taylor, 1975):

$$\mathbf{P}\,(X(t+h) = i+1 \mid X(t) = i) = \lambda_i + o\,(h) \qquad i = 0, 1, \ldots, N-1 \tag{6}$$

$$\mathbf{P}\,(X(t+h) = i-1 \mid X(t) = i) = \mu_i + o\,(h) \qquad i = 1, 2, \ldots, N \tag{7}$$

By convention we define $\lambda_N \equiv \mu_0 \equiv 0$. For compactness of notation let us also define, for $a,\ b = 1,\ 2$: $\bar{m}_a \equiv 1 - m_a$, $\bar{r}_{ab} \equiv 1 - r_{ab}$, $u_{ab} \equiv m_1 r_{ab} + m_2 \bar{r}_{ab}$ and $\bar{u}_{ab} \equiv 1 - u_{ab}$. Then, using (1) and the definition of the Moran process, we calculate:

$$\begin{aligned} \lambda_i = \theta\frac{N-i}{N}\Bigg\{&(1-\rho)\left[m_1\frac{i}{N} + m_2\frac{N-i}{N}\right] + \\ &+\rho\left[u_{11}\left(\frac{i}{N}\right)^2 + 2u_{12}\frac{i}{N}\frac{N-i}{N} + u_{22}\left(\frac{N-i}{N}\right)^2\right]\Bigg\} \end{aligned} \tag{8}$$

$$\begin{aligned} \mu_i = \frac{\theta}{\sigma}\frac{i}{N}\Bigg\{&(1-\rho)\left[\bar{m}_1\frac{i}{N} + \bar{m}_2\frac{N-i}{N}\right] + \\ &+\rho\left[\bar{u}_{11}\left(\frac{i}{N}\right)^2 + 2\bar{u}_{12}\frac{i}{N}\frac{N-i}{N} + \bar{u}_{22}\left(\frac{N-i}{N}\right)^2\right]\Bigg\} \end{aligned} \tag{9}$$

These equations correspond to equations (17) and (18) in (Nowak and Schuster, 1989)[3].

We can also now calculate the (unique) stationary probability distribution p_i, $i = 0, 1, 2, \ldots, N$ of the process (Karlin and Taylor, 1975; Stirzaker, 1994) as follows. Set:

$$\begin{aligned} \pi_0 &= 1 \\ \pi_i &= \frac{\lambda_{i-1}}{\mu_i}\pi_{i-1} \qquad i = 1, 2, \ldots, N \end{aligned} \tag{10}$$

[3](Nowak and Schuster, 1989) use a slightly different version of the Moran birth and death process, perhaps to match the quasi-species formalism more closely. The resulting models are qualitatively similar.

Then we have, for $i = 0, 1, \ldots, N$:

$$p_i = \frac{\pi_i}{\sum_{j=0}^{N} \pi_j} \tag{11}$$

We will be interested in the *mean first passage time* (mfpt) (Karlin and Taylor, 1975) of the process from state i to state j. This may be calculated as follows: let U_i denote the mfpt from state i to state $i+1$ ($i = 0, 1, \ldots, N-1$) and V_i the mfpt from state i to state $i-1$ ($i = 1, 2, \ldots, N$). We then have the recurrence relations:

$$\begin{aligned} U_0 &= \frac{1}{\lambda_0} \\ U_i &= \frac{1}{\lambda_i}(1 + \mu_i U_{i-1}) \qquad i = 1, 2, \ldots, N \end{aligned} \tag{12}$$

and

$$\begin{aligned} V_N &= \frac{1}{\mu_N} \\ V_i &= \frac{1}{\mu_i}(1 + \lambda_i V_{i+1}) \qquad i = 0, 1 \ldots, N-1 \end{aligned} \tag{13}$$

The mfpt from state i to state j for $i < j$ is then given by $U_i + U_{i+1} + \ldots + U_{j-1}$ and for $i > j$ by $V_i + V_{i-1} + \ldots + V_{j+1}$. Note that the mfpt's cannot be expressed solely in terms of the stationary probabilities; knowledge of the actual infinitesimal generators is required.

Finally, to simulate the birth and death process we make use of the following (Karlin and Taylor, 1975): the process waits in the state i for a period of time distributed exponentially with parameter $\lambda_i + \mu_i$. It then makes a transition to state $i+1$ (if $i < N$) with probability $\frac{\lambda_i}{\lambda_i + \mu_i}$, or to state $i-1$ (if $i > 0$) with probability $\frac{\mu_i}{\lambda_i + \mu_i}$.

5 Behaviour of the Model

In the results that follow we have used a short seqeunce length ($\nu = 10$) and population size ($N = 100$) to make the pertinent features of the model clearer. All results extend to higher sequnce lengths and larger populations. Fig. 1 below plots the stationary distribution of the birth and death process for a few values of the mutation rate, all other parameters remaining fixed. We see that at low mutation rates the optimum genotype frequency is generally high; the process spends most of its time with a high proportion of the population "on the spike". It appears unimodal, but there is actually another mode at 0, not visible at this scale. At a slightly higher mutation rate the bimodality becomes more pronounced and the position of the righmost mode shifts to a lower optimum genotype frequency. At a critical mutation rate above this an inflexion point appears and the distribution becomes unimodal. Following (Nowak and Schuster, 1989) we identify this critical mutation rate with the error threshold[4]. Beyond the error threshold the distribution is unimodal and the process spends most of the time with optimum genotype frequency close to zero. Figs. 2, 3 and 4 illustrate the effects of increasing recombination rate on the dynamics of the process in the sub-error threshold regime. In each case the left-hand figure shows the stationary probability distribution while the right-hand figure plots the results of a simulation of the process with the same parameter values. In all figures the (arbitrary) timescale $\theta = 100$, $\sigma = 2$, $\nu = 10$ and $N = 100$. A subtlety in comparing the dynamics for different values of ρ is that changing the

[4]Note that this is not the only possible definition of the error threshold for finite populations. See e.g. (Forst et al., 1995).

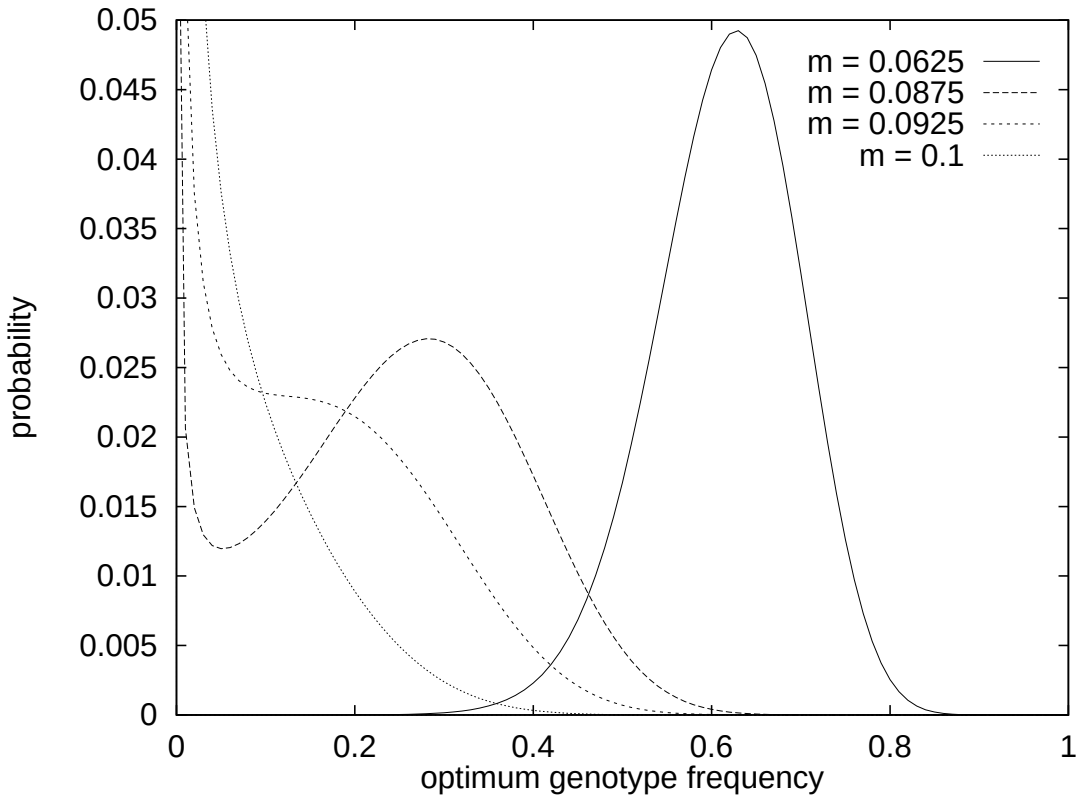

Figure 1: Stationary distribution of the birth and death process for a few values of mutation rate μ. Other parameters are: $\nu = 10$, $N = 100$, $\sigma = 4$ and $\rho = 0.3$.

recombination rate alters the shape of the stationary distribution. Indeed in (Boerlijst et al., 1996) and (Ochoa and Harvey, 1999) it is demonstrated that increasing the recombination rate lowers the error threshold. Thus to establish a baseline for comparison we followed the following procedure: for each value of ρ the mutation rate μ was adjusted so that the *median* of the stationary distribution coincides with the optimum genotype frequency at which the stationary distribution takes on its minimum value between the modes; thus the process spends equal amounts of time in states above and below the optimum genotype frequency dividing the modes.

Figs. 5 and 6 plot the between-mode minimum probability and right-hand mode optimum genotype frequency respectively against recombination rate. Fig. 7 plots the mfpt of the process from the right-hand mode down to the zero state. Again in all figures the mutation rate is adjusted so that the stationary median coincides with the between-modes minimum. In all plots $\theta = 1$, $\sigma = 4$, $\nu = 10$ and $N = 100$.[5]

The effects of increasing the recombination rate now become clear:

1. The time spent by the process in states *between* the modes decreases
2. The optimum genotype frequency of the right-hand mode increases; the modes are "pulled apart"

[5]The "wobbliness" of these plots is due to the fact that there is, for a given recombination rate, a (small) range of mutation rates for which the median equals the between-modes minimum. There was thus some leeway in the precise choice of mutation rate.

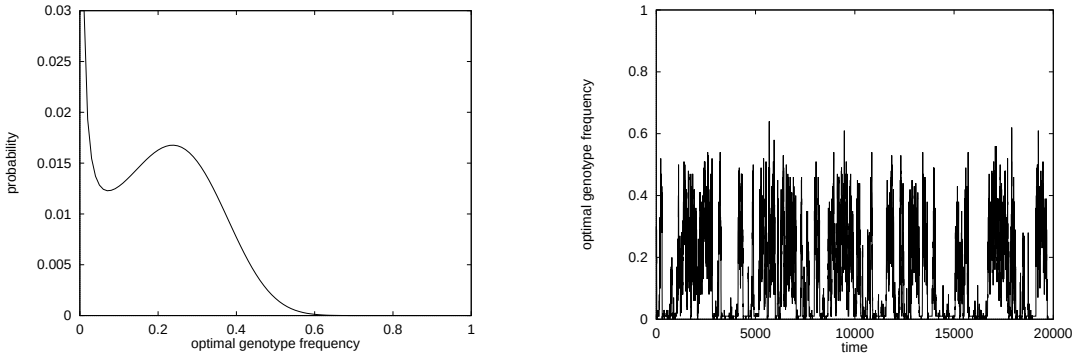

Figure 2: Parameters are: $\mu = 0.0527$, $\rho = 0$.

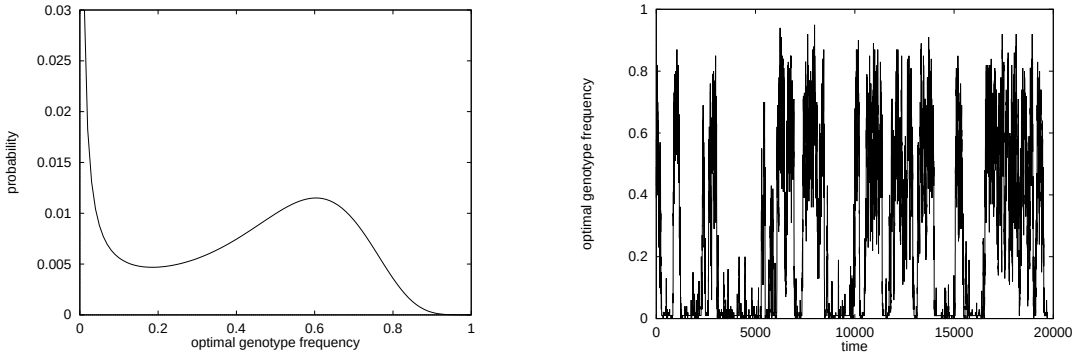

Figure 3: Parameters are: $\mu = 0.01612$, $\rho = 0.45$.

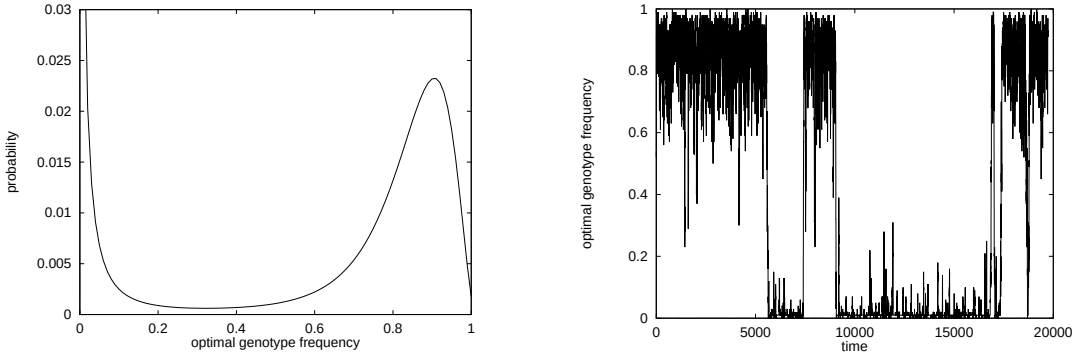

Figure 4: Parameters are: $\mu = 0.005174$, $\rho = 0.6$.

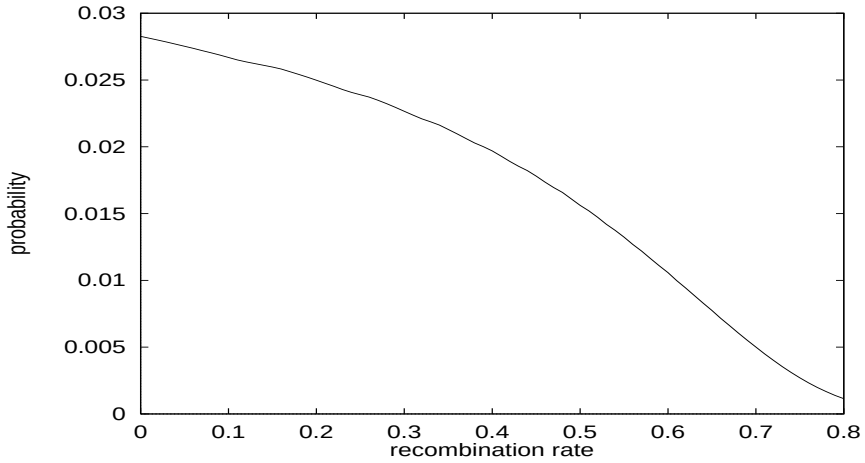

Figure 5: Between-mode minimum probability.

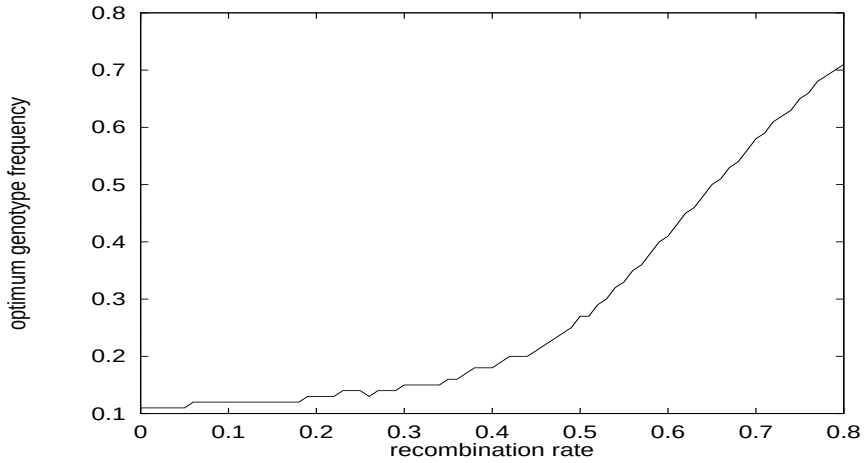

Figure 6: Right-hand mode optimum genotype frequency.

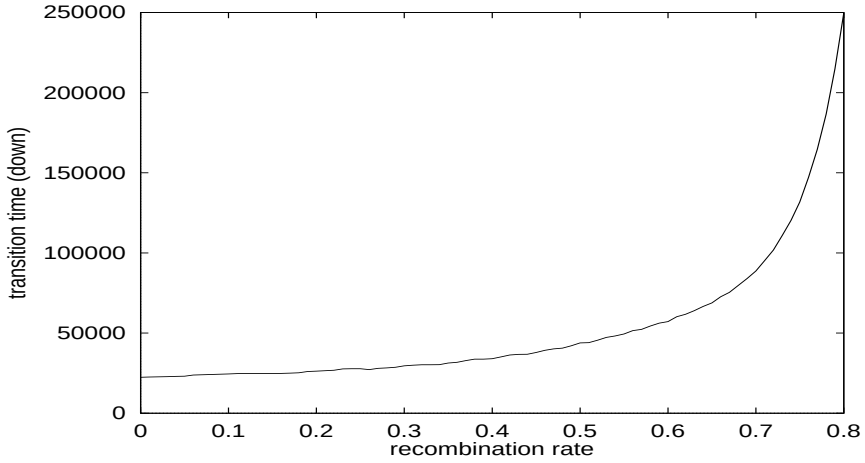

Figure 7: Mean first passage time from right-hand mode down to zero state.

3. The expected waiting time for transitions between the modes "blows up" rapidly

Consider now, for example, the right-hand plot in Fig. 4 and suppose we were to watch the process over a period of time very small compared with the mean between-mode transition time. If we happened to observe the state at a given time to be near one mode it is likely that we should never see it make a transition to the other mode; as far as we could tell the process would be settled in a unimodal steady state. If, however, as is equally likely (recall that by construction the process spends more or less equal amounts of time near each mode), we happened to observe the process near the *other* mode, we would consider the process to be settled in a *different* steady-state. In short, on a timescale small compared to the mean between-modes transition time the process appears bistable. We term this phenomenon *stochastic bistability*.

We see from points 1 and 2 above that at low recombination rates it is more difficult to "separate" the modes - it is thus more difficult to discern a bistable situation (cf. the right-hand plots in Figs. 2 and 3). From point 3 we see that for any given observational timescale we are less and less likely to see a transition between modes as the recombination rate is increased.

6 Conclusions

We have demonstrated that stochastic bistability arises in a finite population as the result of two factors: bimodality of the steady-state distribution and between-mode transition times long compared to the observer's timescale. We have seen that increasing the recombination rate accentuates both of these factors (in the sense of points 1-3 of the previous section). It is of interest to note that, strictly speaking, this form of bistability is present in our model even with *no* recombination present, albeit not readily discernible even at very short timescales due to the poor separation of the modes. We note that for any (Markovian) stochastic evolutionary process which is irreducible (and this would seem to include most finite-population models in population biology) there is a unique stationary distribution. For such processes, therefore, it seems likely that bi- (or multi-) stability must always arise in a similar fashion to our model. Of course there are many stochastic evolutionary scenarios that cannot be modelled by an irreducible Markov process or, indeed, by a Markov process. Nonetheless the phenomenon would appear to be very general.

It would be of great interest to connect the stochastic bistability observed in our model with the bistability observed in infinite-population deterministic models. We speculate that there is a limiting procedure whereby the dynamical equations describing the time evolution of our birth and death model (the forward or backward Chapman-Kolmogorov equations (Karlin and Taylor, 1975)) converge to quasispecies-like differential equations which bifurcate as in the simplified model presented in (Boerlijst et al., 1996).

We also note that, in principle at least, our model allows us to calculate approximations for the error threshold in a finite population where recombination is present, along the lines of (Nowak and Schuster, 1989). In that paper the optima of the stationary distribution are approximated by treating the frequency of optimal genotypes i/N as a continuous variable x. The positions of the optima are then revealed as the solutions of a quadratic equation for x, the discriminant of which (a quadratic in the quantity Q) yields the mutation rate at which the optima coalesce and the distribution becomes unimodal; i.e. the error threshold. In our case recombination introduces a cubic term to the equations, making (analytical) solution more difficult. We hope to carry out a mathematical analysis in a future paper.

Acknowledgements

The author wishes to thank Inman Harvey, Gabriela Ochoa and the Sussex University Neutral Networks group for helpful discussions.

References

Boerlijst, M. C., Bonhoeffer, S., and Nowak, M. A. (1996). Viral quasi-species and recombination. *Proc. R. Soc. Lond. B.*, 263:1577–1584.

Crow, J. F. and Kimura, M. (1970). *An Introduction to Population Genetics Theory.* Harper and Row, New York.

Derrida, B. and Peliti, L. (1991). Evolution in a flat fitness landscape. *Bull. Math. Biol.*, 53(3):355–382.

Eigen, M., McCaskill, J., and Schuster, P. (1989). The molecular quasispecies. *Adv. Chem. Phys.*, 75:149–263.

Forst, C. V., Reidys, C., and Weber, J. (1995). Evolutionary dynamics and optimization: Neutral networks as model-landscapes for RNA secondary-structure folding-landscapes. In Moran, F., Moreno, A., Merelo, J. J., and Chacon, P., editors, *Proc. ECAL '95, Lecture Notes in Artificial Intelligence, vol. 929: Advances in Artificial Life.* Springer Verlag.

Karlin, S. and Taylor, H. M. (1975). *A first course in Stochastic Processes (2nd ed.).* Academic Press (NY).

Kimura, M. (1983). *The Neutral Theory of Molecular Evolution.* Cambridge University Press.

Moran, P. A. P. (1958). The effect of selection in a haploid genetic population. *Proc. Camb. Phil. Soc.*, 54:463–467.

Nowak, M. and Schuster, P. (1989). Error thresholds of replication in finite populations: Mutation frequencies and the onset of Müller's ratchet. *J. Theor. Biol.*, 137:375–395.

Ochoa, G. and Harvey, I. (1999). Recombination and error thresholds in finite populations. In *Proceedings of FOGA 5 (Foundations of Genetic Algorithms).* ftp://ftp.cogs.susx.ac.uk/pub/users/inmanh/fogat.ps.gz.

Stirzaker, D. (1994). *Elementary Probability.* Cambridge University Press.