Consensus Sequence Plots and Error Thresholds: Tools for Visualising the Structure of Fitness Landscapes

Gabriela Ochoa

COGS – The University of Sussex, Falmer, Brighton BN1 9QH, UK
gabro@sussex.ac.uk

Abstract. This paper investigates the occurrence of *error thresholds* in genetic algorithms (GAs) running on a wide range of fitness landscape structures. The error threshold, a notion from molecular evolution, is a critical mutation rate beyond which the evolutionary dynamics of a population changes drastically. The paper also introduces *Consensus sequence* plots, an empirical tool for locating error thresholds on complex landscapes. This plots were borrowed and adapted from theoretical biology. Results suggest that error thresholds occur in GAs but only on landscapes of certain degree of ruggedness or complexity. Moreover, consensus sequence plots can be useful for predicting some features of a landscape such as ruggedness and "step-ness". We argue that error thresholds and consensus sequence plots, may become useful tools for analysing evolutionary algorithms and visualising the structure of fitness landscapes.

1 Introduction

The *error threshold* — a notion from molecular evolution — is the critical mutation rate beyond which structures obtained by a evolutionary process are destroyed more frequently than selection can reproduce them. With mutation rates above this critical value, an optimal solution would not be stable in the population. The main purpose of this paper is to bring the notion of error threshold from molecular evolution to evolutionary computation. In other words, to explore whether a phenomenon similar to error thresholds is found in evolving populations of bit-strings using a GA. In a previous work, we demonstrated empirically the existence of error thresholds on simple abstract landscapes using a standard GA [Ochoa and Harvey, 1998], we found also that the magnitude of error thresholds is lower when recombination is used. This paper, extendeds those findings by studying more complex landscapes, including a real-world engineering problem. A general empirical approach for locating error thresholds on complex landscapes is provided. This approach produces the so-called *consensus sequence* [1] plots. The existence and characteristics of the error threshold on a given landscape depend upon the structure of the landscape. Thus consensus sequence plot can be useful tools for visualising the structure of a given landscape.

[1] The term *sequence* is here interchangeable with string or genotype.

2 Quasispecies and Error Thresholds

This section introduces the notions of *quasispecies* and *error thresholds* from molecular evolution. Quasispecies theory was derived by [Eigen and Schuster, 1979] to describe the dynamics of replicating nucleic acid molecules under the influence of mutation and selection. The theory was originally developed in the context of pre-biotic evolution (studies of the origin of life), but in a wider sense it describes any population of reproducing organisms. A quasispecies is defined as the stationary population distribution of replicating macromolecules under mutation and selection.

The most prominent feature of the quasispecies model is the existence of an error threshold of replication. If replication were error free, no mutants would arise and evolution would stop. On the other hand, evolution would also be impossible if the error rate of replication were too high (only a few mutations may produce an improvement, but most will lead to deterioration). The notion of error threshold allows us to quantify the resulting minimal replication accuracy that still maintains adaptation. The quasispecies model, as stated originally, considered infinite asexual populations. Later extensions included finite populations and recombination. Most quasispecies studies considered simple landscapes, including single peak, double peak, and flat fitness landscapes. In contrast, the work of [Bonhoeffer and Stadler, 1993], described below, studied the evolution of quasispecies on two complex fitness landscapes.

3 Consensus Sequence Plots

The work of [Bonhoeffer and Stadler, 1993] studied the evolution of quasispecies on two correlated fitness landscapes, the Sherrington Kirkpatrick spin glass and the Graph Bipartitioning landscape. The authors described an empirical approach for locating thresholds on complex landscapes. In this paper, this approach is borrowed and adapted. Instead of the quasispecies model, a GA is used as the underlying model of evolution. The resulting method can be applied for identifying error thresholds in GAs running on general complex landscapes. The approach is to calculate and plot the consensus sequence at equilibrium for a range of mutation rates. The consensus sequence in a population is defined as the sequence of predominant symbols (bits) in each position; it is plotted as follows: if the majority of individuals has a '1' or '0' in a position i the field is plotted white or black, respectively. The field is plotted grey if the position is undecided. Figure 1, shows an hypotetical population and calculates its consensus sequence. This plot will correspond to a single line in a consensus sequence plot (See Figure **??**). The Equilibrium State is reached when the proportion of different sequences in the population is stationary. This happens when evolution is simulated for a large enough number of generations. In practice, it is considered that the equilibrium is reached when several parameters of the population (e.g. the maximal and average fitness) reach equilibrium. According to [Bonhoeffer and Stadler, 1993] the error threshold may be approached from *below* or *above*, with both methods producing similar results.

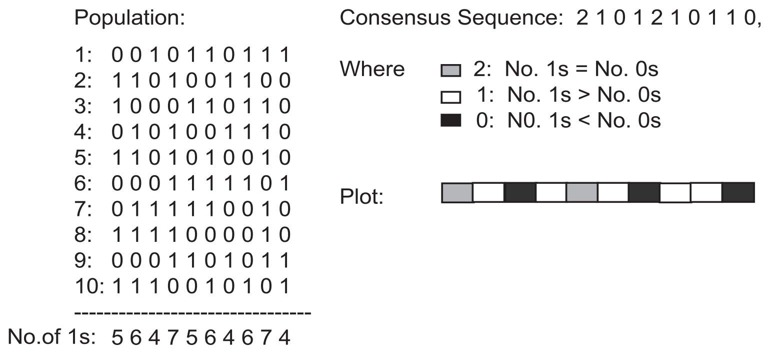

Fig. 1. Calculating and ploting the consensus sequence of a population.

Approaching the Error Threshold from Below

To approach the error threshold from below, the simulation starts with a homogeneous population at the global optimum. This approach requires knowing the optimal string beforehand. Then, the population is allowed to reach equilibrium at a constant mutation rate of $= 0.0$. Afterwards, the mutation rate is increased by a fixed, small step and the computation is continued with the current population. This process is repeated until a predefined maximum for the mutation rate is reached. The plot summarizes a single run, there is no averaging of multiple runs.

Approaching the Error Threshold from Above To approach the error threshold from above, the simulation starts with a random population. Then the population is allowed to reach equilibrium at a constant predefined maximum for the mutation rate [2]. Afterwards, the mutation rate is decreased by a fixed small step and the computation continues with the current population. This process is repeated until the mutation rate is 0.0. Figure **??**. Notice that, in this case, it is not required knowing the optimal string. Hence, in principle, this approach can be used for locating the error threshold on any complex landscape. Again the plot summarizes a single run, there is no averaging of multiple runs.

For both approaches, the consensus sequence in the population is calculated and plotted at the end of each simulation cycle for each mutation step. The error threshold is characterised by the loss of the consensus sequence, i.e. the genetic information of the population. Beyond the error threshold the consensus sequence is no longer constant in time.

---

[2] This value has to be high enough to be above the error threshold for the landscape under study.

```
Procedure Consensus_Plot {
    p = p_max;      /* Initial (high) mutation rate */
    initialise the population (randomly);
    Run_GA;         /* large number of generations */
    Calculate and plot the consensus sequence;
    until (p = 0.0) {
        p = p - p_step;
        Run_GA;     /* large number of generations */
        Calculate and plot the consensus sequence;
    }
}
```

Fig. 2. Algorithm for producing a consensus sequence plot (from above).

4 Test Problems

Two families of abstract fitness landscapes were selected as test problems: the Royal Staircase and the NK landscape. In addition, a real-world engineering problem, the Wing-Box problem, was explored.

Royal Staircase Family of Functions

The *Royal Staircase* family of functions were proposed by [van Nimwegen and Crutchfield, 1998] for analyzing epochal evolutionary search. This functions are related to the previous *Royal Road* functions [Mitchell et al., 1992]. Although simple, Royal Staircase functions capture some essential elements found on complex problems, namely, the existence of highly degenerate genotype-to-phenotype maps (i.e. the mapping from genetic specification to fitness is a many-to-one function). Next, we present a description of the Royal Staircase class of fitness functions:

1. Genotypes are specified by binary strings $s = s_1 s_2 \ldots s_L, s_i \in \{0, 1\}$, of length $L = NK$, where N is the number of blocks and K the bits per block.
2. Starting from the first position, the number $I(s)$ of consecutive 1s in a string is counted.
3. The fitness $f(s)$ of string s with $I(s)$ consecutive ones, followed by a zero, is $f(s) = 1 + \lfloor I(s)/K \rfloor$. The fitness is thus an integer between 1 and $N + 1$, corresponding to 1 plus the number of consecutive fully-set blocks starting from the left.
4. The single global optimum is $s = 1^L$; namely, the string of all 1s.

Fixing N and K defines a particular problem or fitness landscape. For the experiments in this paper we selected two values of N ($N = 3$ and $N = 4$) and a single value of K ($K = 10$), that is, string lengths of 30 and 40 respectively.

NK Family of Landscapes

Kauffman NK family of fitness landscapes [Kauffman, 1989] are determined by two parameters: N and K. N is the string length (of binary strings) and K is the *degree of epistatic interaction* between the bits. An interesting property of the NK landscapes is that the ruggedness of the fitness landscape can be tuned by changing the parameter K. As K increases, the number of optima increases and the fitness correlation decreases.

Wing-Box Problem

The Wing-Box problem was formulated as part of the Genetic Algorithms in Manufacturing Engineering (GAME) [3] project at COGS, University of Sussex. An industrial partner, British Aerospace, provided data from a real Airbus wing box for the definition of the problem. When designing aircraft structures, a common problem is to define structures of minimum weight that can withstand a given load. Figure 3 sketches the elements of a wing relevant to this problem. The wing is supported at regular intervals by slid ribs which run parallel to the aircraft's fuselage. On the upper part of the wing, thin metal panels cover the gap between adjacent ribs. The objective is to find the number of panels and the thickness of each of these panels while minimising the mass of the wing and ensuring that none of the panels buckle under maximum operational stresses. More details, and the equations for calculating the fitness function, can be found in [McIlhagga et al., 1996].

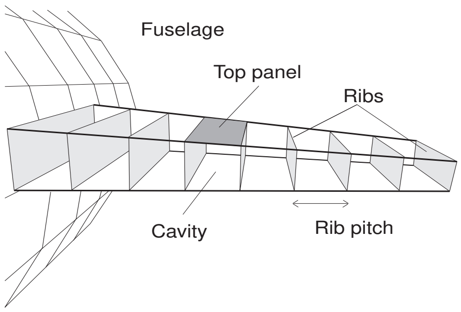

Fig. 3. Relevant elements of a wing. Wing dimensions are fixed. The variable elements are the number of ribs and the thickness of the top panels.

5 Experiments and Results

All experiments were run using a generational GA with fitness proportional selection and a population of 100 individuals. The genetic operations were 2-point crossover with a rate of 0.6 and the standard bit mutation. Mutation rates were

[3] http://www.informatics.sussex.ac.uk/projects/game/

expressed as mutation per genotype. The string length varied according to the landscape under study. The GA was run in two modes: using mutation only (**GA-M**); and using both mutation and recombination (**GA**). Consensus sequence plots were produced for all landscapes under study following the method described in section 3.

5.1 Preliminary Study

A preliminary study was designed to confirm the assertion by [Bonhoeffer and Stadler, 1993] that error thresholds do not depend on whether they are approached from below or from above. Another set of experiments tests whether the error threshold (when approached from above – i.e. from a random population) depends on the particular initial population.

Figure 4 shows the consensus sequence plots on the two selected test problems. In both cases, the error threshold was approached from below and above (Section 3). The Royal Staircase function has a single optimum (the string of all ones). On the other hand, the NK landscape has multiple optima, and they are not know. Consequently, in this case the procedure was to first approach the error threshold from above starting from a random population, then store the consensus sequence thus obtained. Afterwards, the error threshold was approached from below starting from a population where all individuals were copies of the stored consensus sequence. Vertical axes show the mutation rate ranges explored. In all plots, mutation rates are expressed as mutations per bit (m/b). Mutation step-sizes were 0.005 for the Royal Staircase and 0.001 for the NK landscape.

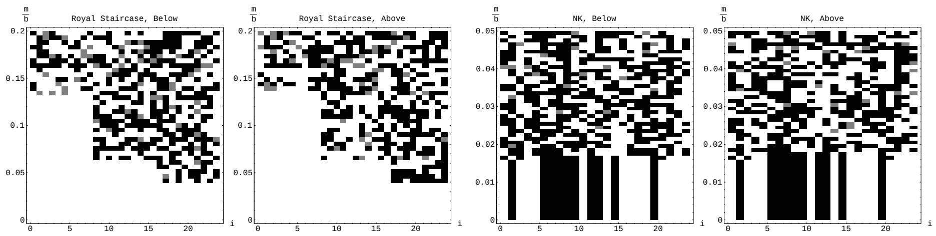

Fig. 4. Consensus sequence plots on a Royal Staircase function ($N = 3$, $K = 10$) and a NK landscape ($N = 24$, $K = 10$). The error threshold is approached from below and above. The horizontal axis show the consensus bit for each position i, the vertical axis show per-bit mutation rates (m/b). Mutation step-sizes were 0.005 (Royal Staircase), and 0.001 (NK). The error threshold is characterised by the loss of the consensus sequence (the string of all 1's for the Royal Staircase; and one local optima for the NK Landscape). For the Royal Staircase, the intermediate error thresholds for each step or fitness level can also be observed.

The plots illustrate the existence of a stable consensus sequence (the string of all 1's for the Royal Staircase; and one local optima for the NK Landscape) for mutation rates below the error threshold. Results confirm that error thresholds do not depend on whether they are approached from below or from above. On the NK landscape the transition occurs close to 0.02 mutations per bit. For the Royal Staircase the critical mutation rate is close to 0.05, in this case, different error thresholds for each fitness level or step can also be observed.

The next set of experiments explores whether error thresholds approached from above (i.e. from a random population) are independent from the initial population. For that purpose, consensus sequence plots are produced for four initial populations (four random seeds) on the same NK landscape (Figure 5).

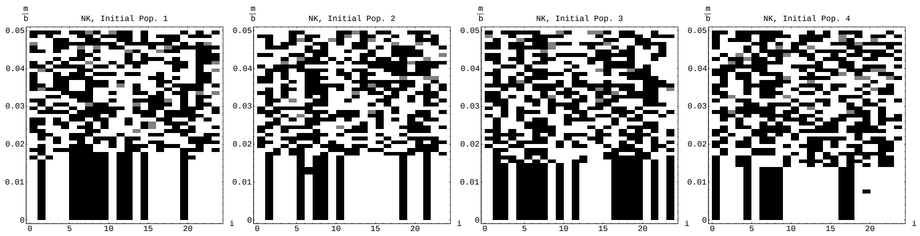

Fig. 5. Consensus sequence plots on a fixed NK landscape ($N = 24$, $K = 10$) for four initial population (different random seeds). The horizontal axis shows the consensus bit for each position i, the vertical axis shows the mutation rate per bit (m/b). The error threshold was approached from above, that is from a random population. Error thresholds are characterised by the loss of the consensus sequence, which is different in each case.

Results suggest that the error threshold is independent of the initial population. Although the consensus sequence achieved in each case is different, the transition occurs at approximately 0.02 mutations per bit in all cases (with a discrepancy of $\simeq 0.002$)

5.2 Rugged Landscapes

Figure 2 shows results for GA (crossover rate = 0.6) and GA-M (crossover rate = 0.0) on a Royal Staircase function with $N = 3$, $K = 10$. The error threshold was approached from below, starting from a homogeneous population at the global optimum and a mutation rate of $p = 0.0$; and from above, starting with a random population and a mutation rate of $5.0/L$ (where L is the string length) . The plots illustrate the existence of a stable consensus sequence for mutation rates below the error threshold. In this case the consensus sequence is the single optimum string in the landscape (the string of all ones). Interestingly, error thresholds do not depend on whether they are approached from below or from above. The error thresholds for each fitness level or step can be clearly observed.

Notice that error thresholds for all fitness levels are lower when recombination is used (GA).

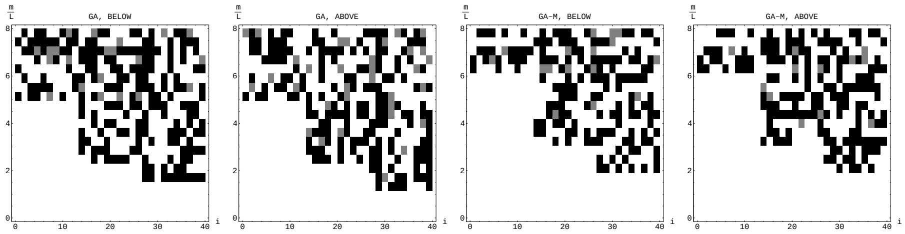

Fig. 6. The Consensus sequence on a Royal Staircase function ($N = 3$, $K = 10$) for GA and GA-M. The horizontal axis show the consensus bit for each position i, the vertical axis show the mutation rate expressed as mutations per genotype. The error threshold was approached from below and above (Section 3). When approached from above, the time of evolution is opposite to the mutation rate axis. The mutation step-size used was $0.25/L$. Error thresholds are characterized by the loss of the consensus sequence, in this case the string of all ones. Also the intermediate error thresholds for each step or fitness level can be observed.

To explore whether increasing the number of blocks would be reflected in the consensus sequence plot, we ran a similar experiment on a Royal Staircase function with $N = 4$, and $K = 10$, for both GA and GA-M (Figure 3). The error threshold was approached from above. The consensus sequence is again the single optimum string in the landscape. The error thresholds for each fitness level or step can be clearly observed. Notice that a fourth level (step) appears compared to Figure 2. Again error thresholds are lower when recombination is used (GA).

Figure 4 illustrates consensus sequence plots for both GA and GA-M on two NK landscapes with $N = 24$; and $K = 6$ (Left), and $K = 12$ (Right). The error threshold is approached from above. Notice that consensus sequences reached by GA and GA-M are different, this is because these landscapes are multimodal, and different runs will most likely lead to different optima. For the more rugged landscape ($K = 12$) the transition is sharper, and the error threshold magnitude is lower when recombination is used (GA). For the less rugged landscape ($K = 6$) the transition seems to be smoother and there is no clear difference between GA and GA-M.

Figure 5 shows results on the Wing-Box problem for GA and GA-M. The plots show the existence of a stable consensus sequence for mutation rates below the error threshold. The error threshold is n visualized as the transition from a stable consensus sequence to a random sequence of bits. Notice that there is not a clear and single transition; from approximately bit 75 to bit 125 the error threshold looks higher than for the rest of the bits. Otherwise the transition seems to occur around 1.5 to 2.0 mutations per genotype. There is, again, no

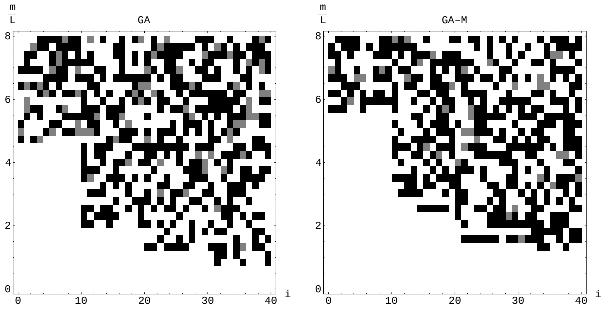

Fig. 7. The consensus sequence on a Royal Staircase function ($N = 4$, $K = 10$), for both GA and GA-M. The horizontal axis show the consensus bit for each position i, the vertical axis show the mutation rate expressed as mutations per genotype. The mutation step-size used was $0.25/L$.

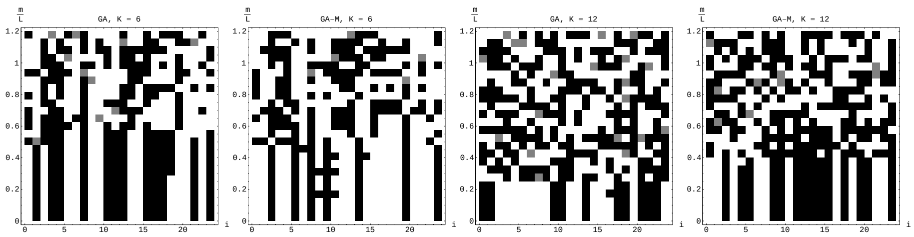

Fig. 8. The consensus sequence on two NK landscapes: $N = 24$, $K = 6$ (Left) and $N = 24$, $K = 12$ (Right). For both GA and GA-M. The mutation step-size used was $0.05/L$. Notice that the error threshold transition is sharper for the more rugged landscape ($K = 12$).

clear difference between GA and GA-M regarding the magnitude of the error threshold.

5.3 Smooth Landscapes

Figure 6 illustrates the consensus sequence on smooth landscapes for both GA and GA-M. Left, NK landscape with $N = 24$ and $K = 2$. Right, Onemax function with $L = 30$. Here there is no clear error threshold, but instead there is a wide band of mutation rates where the consensus sequence is still present in the population.

6 Discussion

The occurrence of error thresholds in evolving populations of bitstrings using a GA (with and without recombination) is verified on several landscapes, includ-

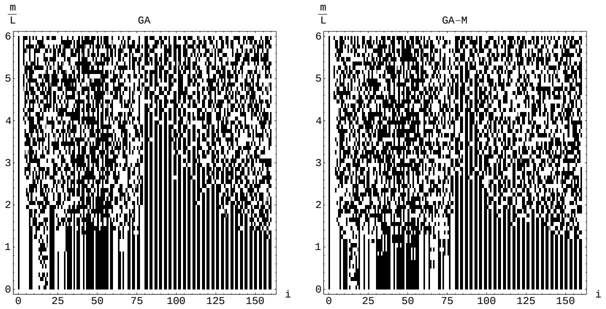

Fig. 9. The consensus sequence on the Wing-Box problem, for both GA and GA-M. The mutation step-size used was $0.1/L$.

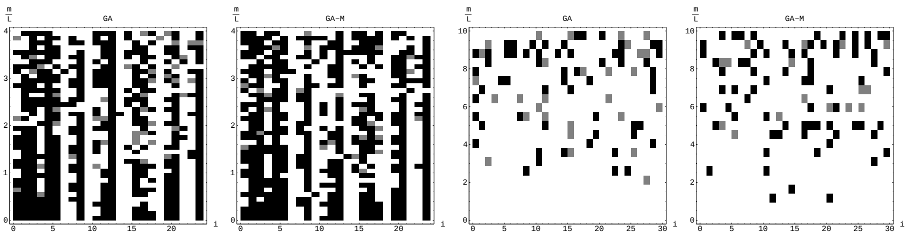

Fig. 10. The consensus sequence on smooth landscapes, for GA and GA-M . Left, NK landscape ($N = 24$, $K = 2$, mutation step $= 0.1/L$). Right, Onemax problem ($L = 30$, mutation step $= 0.5/L$). In these plots there is no clear error threshold, the consensus sequence is still present in the population for high mutation rates.

ing a real-world application. In this way, the notion of error threshold, (already introduced for very simple landscapes [Ochoa and Harvey, 1998]) is brought to evolutionary computation. In previous work, we have shown that error thresholds seems to be related to the more familiar notion of an optimal mutation rate in GAs [Ochoa et al., 1999]. The implication of this finding is two-fold. First, theoretically, in helping us to understand GAs' behavior, as insights about error thresholds will shed light on our understanding of optimal mutation rates. Second, practically, as heuristics for finding error thresholds will provide useful guidelines for setting optimal mutation rates, thus improving the performance of GAs.

The existence of error thresholds depends upon the structure of the fitness landscapes. For smooth landscapes, there is no clear error threshold, the consensus sequence is still present in the population for high mutation rates. For rugged landscapes, on the other hand, there is a clear transition between an "ordered" (selection-dominated) regime and a "disordered" (mutation-dominated) one. For discontinuous (Royal Staircase functions, Figures 2 and 3) and very rugged landscapes (NK with $K = 12$, Figure 4 Right) the error threshold was lower when recombination is used. This effect was not observed for the other landscapes explored in this paper.

This paper also introduced the *consensus sequence plots*. These plots, borrowed and adapted from theoretical biology [Bonhoeffer and Stadler, 1993], are new to evolutionary computation. They represent a novel way to visualize the structure of fitness landscapes, since features such as the "step-ness" of the Royal Staircase function can be clearly noticed (Figures 2 and 3). Moreover, the degree of ruggedness in a landscape was shown to be revealed by the plot (Figures 4 and 6). Consensus sequence plots may also serve as a tool to differentiate critical (and less critical) areas in the genotype, which may have practical implications when tackling real-world problems. First, it may be possible to infer important knowledge about an applied problem. Second, it may be possible to refine the genotype representations and optimal schedules for mutation rates.

Acknowledgements Thanks are due to H. Buxton and I. Harvey for their guidance and support. Thanks also to M. Sordo for critical reading of this document. The author is funded by CONICIT, Venezuela.

References

[Bonhoeffer and Stadler, 1993] Bonhoeffer, S. and Stadler, P. (1993). Error thresholds on correlated fitness landscapes. *J. Theor. Biol.*, 164:359–372.

[Eigen et al., 1988] Eigen, M., McCaskill, J., and Schuster, P. (1988). Molecular quasi-species. *J. Phys. Chem.*, 92:6881–6891.

[Eigen and Schuster, 1979] Eigen, M. and Schuster, P. (1979). *The Hypercycle: A Principle of Natural Self-Organization*. Springer-Verlag.

[Kauffman, 1989] Kauffman, S. (1989). Adaptation on rugged fitness landscapes. In Stein, D., editor, *Lectures in the Sciences of Complexity*, pages 527–618. Addison-Wesley, Reading, MA.

[Maynard Smith, 1970] Maynard Smith, J. (1970). Natural selection and the concept of a protein space. *Nature*, 225:563–564.

[McIlhagga et al., 1996] McIlhagga, M., Husbands, P., and Ives, R. (1996). A comparison of search techniques on a wing-box optimisation problem. *Lecture Notes in Computer Science*, 1141.

[Mitchell et al., 1992] Mitchell, M., Forrest, S., and Holland, J. H. (1992). The Royal Road for genetic algorithms: fitness landscapes and GA performance. In Varela, F. J. and Bourgine, P., editors, *Proceedings of the First European Conference on Artificial Life*. MIT Press, Cambridge, MA.

[Ochoa and Harvey, 1998] Ochoa, G. and Harvey, I. (1998). Recombination and error thresholds in finite populations. In Banzhaf, W. and Reeves, C., editors, *Foundations of Genetic Algorithms*, San Francisco, CA. Morgan Kauffman.

[Ochoa et al., 1999] Ochoa, G., Harvey, I., and Buxton, H. (1999). Error thresholds and their relation to optimal mutation rates. In Floreano, J., Nicoud, D., and Mondada, F., editors, *Proceedings of the Fifth European Conference on Artificial Life*. Springer-Verlag.

[van Nimwegen and Crutchfield, 1998] van Nimwegen, E. and Crutchfield, J. P. (1998). Optimizing epochal evolutionary search: Population-size independent theory. Technical Report Preprint 98-06-046, Santa Fe Institute.