
An Examination of Tunable, Random Search Landscapes

R. E. Smith

Department of
Aerospace Engineering
and Mechanics
The University of Alabama
Tuscaloosa, Alabama 35487
rsmith@ics.uwe.ac.uk

J. E. Smith

The Intelligent Computing
Systems Center
The University of
The West of England
Bristol BS16 1QY
jim@ics.uwe.ac.uk

Abstract

This paper carefully considers random landscapes related to Kauffman's NK model. In particular, it considers a superset of this model (the NKP model) recently suggested in the GA-analytic literature. Landscapes are exhaustively examined for both the distribution of local optima relative to the global optima, and for characteristics that would effect juxtapositional (building block based) search. The later is accomplished through a Walsh-based analysis. The results indicate that K and P have distinct effects on peak distribution, K controlling peak placement, and P affecting relative peak height. The results indicate that P has little effect on a landscape's expected juxtapositional complexity. Moreover, the results suggest that landscapes developed by the NKP procedure are unlikely to have substantial juxtapositional complexity. A Walsh-based procedure that embodies the flavor of the NKP procedure, but can allow for greater expected juxtapositional complexity, is suggested.

1 Introduction

In examining evolutionary computation (EC) systems, it is useful to consider the nature of the search problems to which they are applied. A great deal of the genetic algorithm (GA) literature is directed at this examination. Careful consideration of search problems (or search *landscapes* as they are often called) also has uses in descriptive science. For instance, consider Kauffman's examination of random landscapes with tunable levels of *epistasis* as an explanation of complexity in biological systems (Kauffman, 1993). Since epistasis is a phenomenon that is easily related to the GA search process, Kauffman's NK landscapes are receiving increasing use as test problems for GAs.

This paper examines NK landscapes (and their extensions) as EC test problems. Moreover, it forms relationships between Kauffman-based and Walsh-based (deception-based) landscape

analysis and design. Finally, it suggests new methods for test landscape design, and for understanding of real-world search landscapes.

2 Advantages of NK Landscapes

There are two main advantages to using NK landscapes as EC test problems:

- The landscapes have two tunable parameters, (N and K) that can be used to control characteristics (e.g., landscape size and epistasis, respectively) of landscapes.
- For given values of N and K, one can generate a large number of landscapes, at random.

These advantages allow one to easily generate a large number of EC test problems, and examine EC performance versus the parameters N and K. However, the meaningfulness of such examinations depends on the relationship between these parameters and the difficulty of the landscapes for search in general, and EC search in particular. Kauffman's work provides some indication of the relationship between the parameters and certain aspects of landscape complexity. This complexity may (or may not) indicate search difficulty.

The complete relationship between the NK parameters and the difficulty of the search landscapes they induce is not entirely clear, for several reasons. First, Kauffman's descriptions are primarily concerned with the relationship between epistasis and complexity for purposes of understanding biological systems, rather than understanding generalized search problems. Second, Kauffman does not exhaustively examine any landscapes for the relationship between parameters and general properties. Instead, he only considers the results of randomly restarted hillclimbing on these landscapes. Third, Kauffman only considers aspects of the landscapes that pertain to hillclimbing search. In this paper, these will be called *peak distribution characteristics*. Kauffman does not consider aspects of the landscapes that pertain to juxtapositional search operators, like crossover. In this paper, these will be called *juxtapositional characteristics*.

Moreover, it has been recently shown that the NK landscapes only cover a small region of the space of problems with size N and epistasis K. Heckendorn and Whitley (1997) demonstrate this fact by examining limitations on the number of non-zero Walsh coefficients induced by the NK landscape generation process. Heckendorn and Whitley's work, along with earlier work by Altenberg (1994, 1997), suggests a broader class of random landscapes that cover the complete space of size N problems with epistasis K, through the introduction of another parameter, P. We will call these *NKP landscapes*.

This paper provides an extensive analysis of NKP landscapes. The NK landscapes are a subset of the NKP landscapes, and, therefore, this implies a more extensive analysis of NK landscapes as well. The paper attempts to answer several questions:

- Does the P parameter change the complexity characteristics of the NK landscapes examined by Kauffman? If so, how?
- Does an *exhaustive* examination of several NKP landscapes yield complexity characteristics that agree with those reported by Kauffman?
- How does the difficulty of NKP landscapes for *juxtapositional* operators vary with N, K, and P (e.g., how are N, K, and P related to *deception* and other GA-analytic measures of difficulty)?

The practical implications of these examinations are two-fold:

- They provide greater understanding of the expected difficulty of EC test problems generated via the NKP (and therefore, the NK) landscape procedure.

- They suggest possible alternative methods for generating random landscapes with tunable difficulty for use as EC test problems.

Moreover, these examinations may provide further insight into the nature of the search problems induced by real-world problems..

The remainder of this paper begins with an introduction to the NK landscapes, then introduces an analogous set of NKP landscapes, and landscapes based on Walsh functions. This also allows for introduction of the NKP landscape concepts presented by Heckendorn and Whitley. Given this introduction, a detailed examination of the peak distribution characteristics of NKP landscapes is presented. This is followed by a Walsh analysis perspective on the juxtapositional characteristics of NKP landscapes. Final comments include suggestions for alternative methods of generating random, tunable landscapes.

3 Kauffman's NK Landscapes

In this discussion we will assume all genes are binary, for convenience. However, the results are extensible to problems with larger gene alphabets.

Specifying an NK landscape requires the following parameters:

N - the total number of bits (genes).

K - the amount of epistasis. Each bit depends on K other bits to determine that its fitness contribution. We will call the K+1 bits involved in each contribution a *subfunction*.

b_i - N (possibly random) bitmasks ($i = 1, 2, 3, \dots, N$). Each bitmask is of length N, and contains K+1 ones. The 1s in a bitmask indicate the bits in an individual that are used to determine the value of the i th sub-function.

Given these parameters, one can construct a random NK landscapes as follows:

- Construct an N by 2^{K+1} table, **X**.
- Fill **X** with random numbers, typically from a standard uniform distribution.

Given the table **X**, and the bitmasks, one determines the fitness of an individual as follows:

- For each bitmask b_i , select out the substring of the individual that correspond with the K+1 one-valued bits in that bitmask.
- Decode these bits into their decimal integer equivalent j .
- Add the entry $A(i, j)$ to the overall fitness function value for this individual.

Note that the fitness values are normalized by dividing by N.

A typical set of bitmasks for this type of problem consists of all N bitmasks that have K+1 consecutive ones. In this case the string is treated as a circle, so that the consecutive 1 bits wrap around. This set of bitmasks outlines a function where any given bit depends on the K preceding bits to determine its contribution to fitness. However, bitmasks are sometimes used such that b_i has the i th bit set to one, but the remaining K one-valued bits are selected at random. Some other possibilities are discussed in Altenberg (1997).

4 NKP Landscapes

Altenberg (1994, 1997), and later Heckendorn and Whitley (1997), allude to a set of landscapes that we will call the NKP landscapes. If one uses an analogy to the NK landscapes, specifying an NKP landscape requires the same parameters as the NK landscapes, with one addition:

P - the number of sub-functions (each assumed to be of epistasis K , for this discussion) that contribute to the overall fitness function. This discussion also assumes that the P sub-functions are simply summed to determine overall fitness. Each sub-function is associated with a bitmask (also called a *partition*) b_i . Note that, for coverage of the space of all K epistatic functions

$$0 \leq P \leq \binom{N}{(K+1)}.$$

Moreover, this means that one must specify P bitmasks, rather than only N . This means that table **X** must be extended to be P by 2^{K+1} . Otherwise, the construction and decoding procedure is the same as that for NK landscapes.

5 Walsh-based Epistatic Landscapes

A set of functions similar to the NKP landscapes can be constructed using Walsh coefficients. This is convenient, since Walsh coefficients have often been used in GA analysis. Moreover, it will help to illustrate Heckendorn and Whitley's conclusions. However, one must make note of what is meant by epistasis in Walsh-based landscapes. Some terminology will be useful here.

Consider a search space defined over length L binary strings. In terms of the Walsh coefficients, define the fitness of an individual as

$$f(x) = \sum_{j=0}^{2^L} w_j \prod_{i=1}^L \Psi(x_i, J_i)$$

where x is the bit string representing the individual, x_i is the i th bit in that string, w_j is the Walsh coefficient corresponding to the *partition* (bitmask) numbered j , J is the bitmask (which is the N -bit binary integer representation of j), J_i is the i th bit in the bitmask, and Ψ is the function describe in Table 1:

x_i	J_i	$\Psi(x_i, J_i)$
0	0	1
0	1	1
1	0	1
1	1	-1

Table 1: Walsh Transform Function Ψ

Let us call the *order* of a Walsh coefficient the number of ones in the partition (bitmask) associated with that coefficient. If a function has a non-zero Walsh coefficient of a given order (say, $K+1$), that means that there is a fitness contribution that depends on all $K+1$ bits in the associated partition.

Given this definition, one can say that the epistasis (in the sense of Kauffman) of a Walsh-based function is the maximum order of the non-zero Walsh coefficients, minus one.

If one uses an analogy to the NKP landscapes, specifying a similar type of Walsh-based landscape is straight forward. It requires the same set of parameters as the NKP landscapes. Given these parameters, one can construct a random Walsh-based function of length N and epistasis K as follows:

- A.** Construct a P by 2^{K+1} table $\mathbf{W}$.
- B.** Fill the table $\mathbf{W}$ with random numbers, taken from some distribution, possibly a normal distribution.

One determines the fitness of an individual as follows:

- A.** For each bitmask b_i , select out the bits in the individual that correspond with the $K+1$ 1-valued bits in the bitmask. Call this length $K+1$ bitstring S_i .
- B.** For each partition order m , $m=0, \dots, (K+1)$, determine the parity of all the order m partitions of S_i . That is, determine the parity of each subset of m bits that can be taken from string S_i . Note that there are a total of 2^{K+1} partitions that must be examined for parity. Each of these corresponds to a column in the random table $\mathbf{W}$ described above.
- C.** If the parity of the j th partition of S_i is even, add table entry $\mathbf{W}(i, j)$ to the overall fitness. If the parity of the j th partition of S_i is odd, subtract table entry $\mathbf{W}(i, j)$ from the overall fitness.

Note that if the bitmasks b_i are non-overlapping, the values in the random table correspond directly to the Walsh coefficients for various, associate partitions in the overall fitness function. However, if the bitmasks overlap, the direct correspondence to the overall Walsh coefficients is less clear. However, as Heckendorn and Whitley (1997) point out, the overlap in the bitmasks affects only the Walsh coefficients associated with partitions in the overlapping region. The Walsh coefficients in these overlapping regions must have an order less than $K+1$.

The relative scale of these landscapes can be changed without loss of generality. Given this consideration of arbitrary scale, the Walsh based procedure and the NKP procedure outlined above can both generate all landscape “shapes” with epistasis K .

Clearly, the NK landscape procedure can only generate a subset of the landscapes generated by the Walsh or NKP procedure. Therefore, the NK landscapes can only generate a subset of the space of landscapes with epistasis K . For a more detailed proof this, see Heckendorn and Whitley (1997).

6 What Are The Effects of P ?

We have noted that NKP landscapes (neglecting scale) cover the space of possible problems with K epistasis. We have also noted that NK landscapes only cover a subset of this space. Given this, one must ask whether the NK landscapes are indicative of the general behavior of the larger space of NKP landscapes. In other words, what are the effects of the new parameter P on the complexity of these tunable, random landscapes. If the effect is negligible on the complexity character of the landscapes, one might not need to consider P in the design of random landscapes as indicative test problems.

This section will present an examination of the effects of N , K , and P on various aspects of problem complexity. Part of the examination will specifically focus on the peak distribution

characteristics considered by Kauffman for NK landscapes. Since the NK landscapes are a subset of the NKP landscapes, this will allow a reconsideration of the results presented by Kauffman. However, the examination presented here will differ from Kauffman's, in that the landscapes considered will be examined exhaustively. Kauffman only examined landscapes via repeated, random restarts of a hillclimbing algorithm.

7 Peak Distribution Characteristics

In order to evaluate the properties of these NKP landscapes, a number of landscapes were created for each of a set of combinations of parameter values. For landscapes of size $N = 12, 16$ and 20 , 10 landscapes were created for each of the combinations with K varying between 2 and 8 , and P taking the values $\{4, 8, 16, 32, 64\}$. Each landscape was exhaustively probed in order to discover the location and value of all the optima (both local and global), and the following values were then calculated for each landscape:

- Maximum, minimum, mean and standard deviation of fitness,
- Mean and standard deviation of optima fitness,
- Number of optima, mean and standard deviation of distance of optima from the global optimum,
- Correlation between relative fitness of local optimum and distance from global optimum.

The results were subjected to a one-way *analysis of variance* (ANOVA) for the factors N , K and P , and to tests for linearity. The results are summarized in Table 2. Where the ANOVA tests indicate statistical significance at the 1% level, the ϵ^2 value (a measure of the proportion of the measured value that can be attributed to the factor) is given (and particularly meaningful values are shown in bold). Where the linearity tests indicate statistical significance at the 1% level the correlation coefficient in the linear model is given. In all cases, NSS indicates no statistical significance

Landscape	N		K		P	
Measure	anova	linearity	anova	linearity	anova	linearity
Land Max.	0.029	0.170	0.042	0.178	0.906	-0.948
Land Min.	0.035	-0.187	0.041	-0.178	0.894	0.942
Land Avg.	NSS	NSS	NSS	NSS	NSS	NSS
Land Std. Dev.	NSS	NSS	0.020	0.117	0.988	-0.973
Number of Peaks	0.394	0.587	0.292	0.519	NSS	NSS
Mean Distance	0.841	0.917	0.098	0.249	NSS	NSS
Std. Dev. Distance	0.339	0.582	0.366	-0.572	NSS	NSS
Optima Correlation	0.011	0.102	0.641	0.765	NSS	NSS
Mean Optima Fit.	0.025	0.158	NSS	NSS	0.898	-0.935
Std.Dev. Optima Fit.	NSS	-0.087	0.041	0.155	0.816	-0.900

Table 2: Statistical factors indicating the effects of N , K , and P on peak distribution characteristics of landscapes. The Optima Correlation row pertains to the correlation between local optima fitness, and distance from the global optima.

Inspection of the results above shows that the variables K and P affect the nature of the landscape in completely different ways, which do not interact. In many cases, even when a factor's effect is a statistical significant, the proportion of variation that can be attributed to that factor is minor. An example of this is in the value of the landscape maximum, where both N and K have a statistically significant effect, but 90.58% of the variation can be attributed to the effect of P .

The first thing to note is that the landscape mean fitness is unaffected by either K or P. The mean is always 0.5 (within experimental error). Next, note that the P parameter effectively acts as a kind of “gain control” on the fitness values’ deviation from this mean. As the number of partitions considered (P) increases, the values of the maxima decrease towards 0.5, and those of the minima increase towards 0.5. This effect arises from the compromises enforced by considering increasing numbers of overlapping partitions. This explains the negative correlation between the value of P and the mean optima fitness, and between P and the standard deviations of the fitness values of the optima and the landscape as a whole. This is the *complexity catastrophe* noted by Kauffman. However, our results show that it is a result of the number of partitions considered (P) rather than their size (K).

By contrast, the epistasis parameter K has little effect on the global fitness statistics, and does not affect the mean fitness values of the local optima at all. However, along with N, K is responsible for determining the number and dispersion of suboptima in the landscape. In particular, K is largely responsible for the correlation between the fitness (relative to the global optimum) of a local optima, and its Hamming distance from the global optimum. Also note that the number of peaks increases linearly with K.

Kauffman reports (Tables 2.1 and 2.2 in Chapter 2 of Kauffman, 1993) that for $K > 0$, the mean fitness of optima found by hill-climbing decreases with increasing K. In fact, study of the results quoted in those tables shows that the mean values are mostly within one standard deviation of each other. Analysis of the exhaustive searches done above shows that there is not a statistically significant relationship between the size of the partitions considered and the fitness range of the landscapes. This can be explained by considering the effect of changing the allele value at a single position, whilst leaving the rest of the string unchanged. The way in which the fitness is calculated means that for each partition in which the position is included, one of two random values will be taken according to the allele value. Thus if a bit position is included in x partitions, then on average x/2 of them will take the lesser of the two possible values.

Assuming the bit masks are randomly created, the probability that the bit position is included in any one partition is simply $(K+1)/N$, and by extension the probability that the position is included in exactly x partitions is

$$\text{Prob(in x partitions)} = \binom{P}{x} \left(\frac{k+1}{N} \right)^x \left(1 - \frac{k+1}{N} \right)^{P-x}$$

Hence $E(x)$, the expected number of partitions that a given bit position is included in is

$$\begin{aligned} E(x) &= \sum_{x=0}^P x \binom{P}{x} \left(\frac{k+1}{N} \right)^x \left(1 - \frac{k+1}{N} \right)^{P-x} \\ &= \sum_{x=0}^P \frac{P}{N} \binom{P-1}{x-1} (k+1)^x (N-k-1)^{P-x} \end{aligned}$$

This shows that even considering the allele value in a single position, the number of partitions which can be expected to take sub-optimal values scales as an exponential of P, and only as a polynomial of K. Further, for fixed P, as the value of K changes two rightmost terms (i.e., those involving K) change value in opposing directions, counteracting each other.

Although the epistasis parameter K does not affect the scale of fitness values, it, rather than P , is responsible for determining the number and dispersion of sub optima in the landscape. As K increases so the number of peaks increases linearly.

Kauffman (1993, Figures 2.7 and 2.8 in Chapter Two) shows local optima as points on plots of relative fitness versus hamming distance from the global optima. This yields “clouds” of points that can be characterized by the associated N , K , and P values. The mean distance corresponds to the midpoint of the cloud on the distance axis and the standard deviation is a measure of the “width” of the cloud. The correlation coefficient is the tangent of the angle between the principal axis of the cloud and a line drawn through the midpoint, parallel to the fitness axis. Specifically, a correlation coefficient of zero indicates that the cloud is symmetrical about the midpoint to the fitness axis. We summarize measures similar to those presented in Kauffman’s cloud diagrams in Figure 1. The plotted values are for $N = 16$, and are averaged over P .

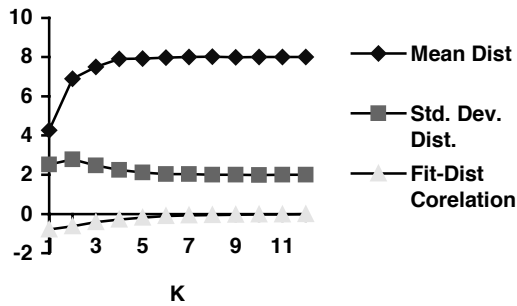


Figure 1: Variation of peak distribution characteristics versus K for $N = 16$ (averaged over P).

Broadly speaking, there are two regimes of behavior in this graph. For K less than or equal to approximately 2, the peaks are on average close to the global optima, and there is a negative correlation between relative fitness and distance from the optimum. This is the “Massif Central” effect noted by Kauffman and confirmed by Hordijk & Manderick (1995) and Smith and Fogarty (1996).

However, for K greater than 3, the peaks appear to be randomly spaced. The mean distance from the optimum is $N/2$. The standard deviation in the distance drops from around 2.5 ($K=1,2,3$) to 2.0, and the correlation between fitness and distance from the optimum drops from -0.79 ($K=1$) to $K = -0.109$ ($K = 6$).

8 Juxtapositional Characteristics

To consider how N , K , and P affect juxtapositional characteristics of the landscape (that is, patterns in the landscape that will affect the utility of juxtaposing building blocks with recombination operators), some elaboration is necessary. Define the global optima in a landscape as the value of bitstring x that yields the highest value of $f(x)$. Call this optimal string x^* :

$$f(x^*) \geq f(x), \text{ for all } x.$$

This condition for optimality, in terms of ordered Walsh coefficients is:

$$\sum_{m=0}^L \sum_{o(w_j)=m} w_j \prod_{i=1}^L \Psi(x_i^*, J_i) \geq \sum_{m=0}^L \sum_{o(w_j)=m} w_j \prod_{i=1}^L \Psi(x_i, J_i), \text{ for all } x.$$

Define the n th order fitness of a combination of bits as:

$$f^n(x) = \sum_{m=0}^n \sum_{o(w_j)=m} w_j \prod_{i=1}^L \Psi(x_i, J_i)$$

In other words, this is the fitness of the bit string when one considers only Walsh coefficients up to order n . Note that this corresponds to considering epistatic effects up to order $n-1$. The n th order fitness is related to the effectiveness of juxtapositional operators, since for some value of n , these operators disrupt the associated building blocks. Thus, a search scheme using such an operator will be insensitive to effects of Walsh coefficients of some order greater than n . This examination only considers order, and not defining length. However, analogous arguments can be made for juxtapositional operators that are affected by defining length.

Define the n th order optima as the string x for which $f^n(x)$ is highest. Call this n th order optimal string x^{n*} .

$$f^n(x^{n*}) \geq f^n(x), \text{ for all } x.$$

Thus:

$$\sum_{m=0}^n \sum_{o(w_j)=m} w_j \prod_{i=1}^L \Psi(x_i^{n*}, J_i) \geq \sum_{m=0}^n \sum_{o(w_j)=m} w_j \prod_{i=1}^L \Psi(x_i, J_i), \text{ for all } x.$$

Note Deb's (1997) definition of deception: "A schema partition is said to be deceptive if the schema containing the deceptive optimal solution is no worse than all other schemata in the partition." One can state the inequalities that define deception (Deb, 1997) in terms of the previous definitions, and the associated Walsh coefficients. However, this yields a large number of such inequalities (on the order of 2^L). Therefore, calculating probabilities of strict deception from these inequalities is difficult. To yield a more calculable measure of juxtapositional complexity in random landscapes, we will say that a landscape is n th order misleading if and only if

$$f^n\left(\left(x^*\right)'\right) \geq f^n(x^*).$$

That is, the complement of the optima (roughly corresponding to Deb's "deceptive optima") has an n th order fitness greater than or equal to that of the optima. For this condition to hold:

$$\sum_{m=0}^n \sum_{o(w_j)=m} w_j \prod_{i=1}^L \Psi\left(\left(x^*\right)'_i, J_i\right) - \sum_{m=0}^n \sum_{o(w_j)=m} w_j \prod_{i=1}^L \Psi(x_i^*, J_i) \geq 0.$$

Given the relationship between the Walsh sums of complementary strings, this condition can be rewritten as:

$$\sum_{m=0}^n (-1)^m \sum_{o(w_j)=m} w_j \prod_{i=1}^L \Psi(x_i^*, J_i) - \sum_{m=0}^n \sum_{o(w_j)=m} w_j \prod_{i=1}^L \Psi(x_i^*, J_i) \geq 0.$$

Canceling terms gives:

$$- \sum_{m=0}^{(n-1)/2} \sum_{o(w_j)=(2m+1)} w_j \prod_{i=1}^L \Psi(x_i^*, J_i) \geq 0.$$

Assume that all Walsh coefficients for a given class of binary encoded search spaces are random, with known distribution functions. Further assume:

- that Walsh coefficients of the same order come from the same distribution functions, and
- that pertinent statistics for all the Walsh coefficient distribution functions are known.

Let's assume that it's equally likely that any given string x is x^* . Note that this implies an assumption that the Walsh coefficients (with order greater than zero) have zero mean.

Clearly, the event that some string x is the optima is mutually exclusive from the event that any other string is the optima. Therefore, without loss of generality, let's consider the case where x^* is the length L string of all 0s, which we will call x^0 . This string has the property that it has even parity in all subpartitions. Thus, if $x^{n*}=x^0$, the misleadingness condition is:

$$\sum_{m=0}^{(n-1)/2} \sum_{o(w_j)=(2m+1)} w_j \leq 0$$

Given that $x^*=x^0$:

$$f\left(\left(x^0\right)'\right) - f\left(x^0\right) \leq 0$$

(from the optimality condition) yields:

$$\sum_{k=0}^L (-1)^k \sum_{o(w_j)=k} w_j - \sum_{k=0}^L \sum_{o(w_j)=k} w_j \leq 0 \quad \text{or} \quad \sum_{k=0}^{(L-1)/2} \sum_{o(w_j)=(2k+1)} w_j \geq 0$$

Splitting this into two terms yields:

$$\sum_{k=n/2}^{(L-1)/2} \sum_{o(w_j)=(2k+1)} w_j \geq - \sum_{k=0}^{(n-1)/2} \sum_{o(w_j)=(2k+1)} w_j$$

Given these assumptions and developments, the probability of the misleadingness condition being true, given that $x^{n*}=x^0$ is:

$$P_m\left(x^{n*} = x^0\right) = P\left(\sum_{k=0}^{(n-1)/2} \sum_{o(w_j)=(2k+1)} w_j \leq 0 \mid \sum_{k=n/2}^{(L-1)/2} \sum_{o(w_j)=(2k+1)} w_j \geq - \sum_{k=0}^{(n-1)/2} \sum_{o(w_j)=(2k+1)} w_j\right)$$

Then, the total probability of having a misleadingness will be:

$$P_m = \sum_{x^a} P\left(m \left| \left(x^{n^*} = x^a \right) \right| \right) P\left(x^{n^*} = x^a \right)$$

where the sum is taken over all possible length L strings x^a . There are 2^L equal terms in this sum, therefore, given previous assumptions:

$$P_m = P\left(\sum_{k=0}^{(n-1)/2} \sum_{o(w_j)=(2k+1)} w_j \leq 0 \left| \sum_{k=n/2}^{(L-1)/2} \sum_{o(w_j)=(2k+1)} w_j \geq - \sum_{k=0}^{(n-1)/2} \sum_{o(w_j)=(2k+1)} w_j \right. \right)$$

To summarize: the number given by expression, which we will call the probability of n^{th} order misleadingness, is the probability that, when one considers all building blocks of n bits or less, the compliment of the optima appears as fit or more fit than the optima itself.

Assume the Walsh coefficients (of order greater than zero) are normally distributed and zero mean. Several properties of the normal distribution are applicable. Specifically, recall that the distribution of a sum of normally distributed random variables is itself normal, with mean equal to the sum of the means, and variance equal to the sum of the variances.

Assume that the order k Walsh coefficients are drawn from a normal distribution with mean μ_k and variance σ_k^2 . Then, if we define two new variables A and B :

$$P_m = P(A \leq 0 | B \geq -A)$$

where A and B are both normally distributed zero-mean random variables, with:

$$\sigma_A^2 = \sum_{k=n/2}^{(L-1)/2} \binom{L}{(2k+1)} \sigma_{(2k+1)}^2, \quad \sigma_B^2 = \sum_{k=0}^{(n-1)/2} \binom{L}{(2k+1)} \sigma_{(2k+1)}^2.$$

Assume, for the time being, that A and B are uncorrelated. Given this, the necessary probability has a bivariate normal distribution with correlation coefficient $\rho=0$. However, since we are given that $B > 0$, we are considering only half of this distribution. A bivariate normal (with zero correlation) is given by

$$P(D) = \frac{1}{2\pi\sigma_A\sigma_B} \iint_D \exp\left(-\frac{A^2}{2\sigma_A^2} - \frac{B^2}{2\sigma_B^2}\right) dA dB$$

To evaluate the integral, one can make the transformation (Maghsoodloo & Huang, unpublished):

$$x = A \frac{\sigma_B}{\sigma_A}, dx = \frac{\sigma_B}{\sigma_A} dA.$$

and change to polar coordinates such that:

$$x = r \cos \theta, B = r \sin \theta$$

Then one can construct the appropriate limits of integration to find the desired probability, giving:

$$P_m = \frac{1}{\pi} \int_{\pi - \tan^{-1}\left(\frac{\sigma_A}{\sigma_B}\right)}^{\pi} \int_0^{\infty} \frac{r}{\sigma_B^2} \exp\left(\frac{-r^2}{2\sigma_B^2}\right) dr d\theta$$

Which simplifies to:

$$P_m = \frac{1}{\pi} \tan^{-1}\left(\frac{\sigma_A}{\sigma_B}\right)$$

Note that the maximum value of P_m in this case is 0.5.

Now consider the more general case where A and B are correlated, with correlation coefficient ρ . The probability is still given by a bivariate normal, but this time in its more general form:

$$P(D) = \frac{1}{2\pi\sigma_A\sigma_B\sqrt{1-\rho^2}} \iint_D \exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{A^2}{\sigma_A^2} + 2\rho\frac{AB}{\sigma_A\sigma_B} + \frac{B^2}{\sigma_B^2}\right]\right) dA dB$$

With a rotation transformation, this can be reduced to:

$$P(D) = \frac{1}{2\pi\sqrt{\lambda_1\lambda_2}} \iint_D \exp\left(-\frac{x_1^2}{2\lambda_1} - \frac{x_2^2}{2\lambda_2}\right) dx_1 dx_2$$

where

$$x_1 = A \cos \psi + B \sin \psi,$$

$$x_2 = -A \sin \psi + B \cos \psi,$$

$$\psi = \tan^{-1}\left(\frac{\lambda_1 - \sigma_A^2}{\rho\sigma_A\sigma_B}\right)$$

(rotation clockwise), and λ_1, λ_2 are the eigenvalues of the covariance matrix for A and B , given by:

$$\lambda_1, \lambda_2 = \frac{1}{2} \left[(\sigma_A^2 + \sigma_B^2) \pm \sqrt{\sigma_A^4 + 2(2\rho^2 - 1)\sigma_A^2\sigma_B^2 + \sigma_B^4} \right]$$

Given previous developments, and the appropriate changes in bounds of integration associated with the rotation transformation:

$$P_m = \frac{1}{\pi} \int_{\frac{\pi}{2} + \tan^{-1}\left(\sqrt{\frac{\lambda_1}{\lambda_2}} \cot(\psi)\right)}^{\frac{\pi}{2} + \tan^{-1}\left(\sqrt{\frac{\lambda_1}{\lambda_2}} \cot\left(\psi + \frac{\pi}{4}\right)\right)} \int_0^{\infty} \frac{r}{\lambda_2} \exp\left(-\frac{r^2}{2\lambda_2}\right) dr d\theta$$

giving:

$$P_m = \frac{1}{\pi} \left[\tan^{-1} \left(\sqrt{\frac{\lambda_1}{\lambda_2}} \left(\frac{\rho \sigma_A \sigma_B}{(\lambda_1 - \sigma_A^2)} \right) \right) - \tan^{-1} \left(\sqrt{\frac{\lambda_1}{\lambda_2}} \left(\frac{\rho \sigma_A \sigma_B - (\lambda_1 - \sigma_A^2)}{\rho \sigma_A \sigma_B + (\lambda_1 - \sigma_A^2)} \right) \right) \right]$$

Note that, regardless of the selection of which eigenvalue is which, the numerator in the first inverse tangent term goes to zero faster than the denominator as ρ approaches 0. Therefore, this expression reduces to the previously stated expression for the uncorrelated case when $\rho = 0$, as expected.

In the NKP landscapes we have examined thus far, $\rho \approx 0$. Recall that this case gives a maximum probability of 0.5. This illustrates that while misleadingness is possible in landscapes generated by the NKP procedure, it occurs at most half the time.

9 Evaluating Juxtapositional Complexity

To examine the juxtapositional complexity characteristics of NKP landscapes, a number of landscapes were created for various combinations of N, K, and P values. Probabilities of misleadingness were calculated from these landscapes. Before examining the results, we make some note of the validity of previous assumptions. First, note that, given the random determination of bit masks in the NKP procedures, one can assume that all partitions of similar order should have similar Walsh statistics (over the space of randomly generated landscapes). Second, note that each Walsh coefficient is a sum of 2^L signed fitness values, and that, in a NKP landscape, each of the fitness values is itself a sum of uniform random variables. Therefore, it is reasonable to assume that the statistics of samples from these sums large numbers of uniform random variables should be normally distributed. Under the NKP procedures, these distributions all have zero mean, since all points are equally likely to be the optima in any given partition.

Finally, we note that our experiments indicate that the sums A and B are either completely uncorrelated, or virtually so. All experiments give very low values for correlations between A and B. Moreover, increasing sample sizes (i.e., the number of landscapes considered) resulted in increasingly lower correlation. Therefore, the calculations presented are based on the $\rho = 0$ case discussed above. The implications of correlation between A and B are discussed in a later section.

For landscapes of sizes $N = \{8, 16, 20\}$, 40 landscapes were created for each of the possible combinations of values of $K = \{2, 3, 4, 5, 6\}$ and $P = \{2, 4, 8, 16, 32, 64, 128\}$. Given the assumption of $\rho = 0$, these were sufficient sample sizes for the required calculations. Although these are relatively small values of N, it is important to note that this can be thought of as a subproblem of a larger problem.

Figures 2 through 4 present the calculated probabilities of misleadingness, versus order of misleadingness.

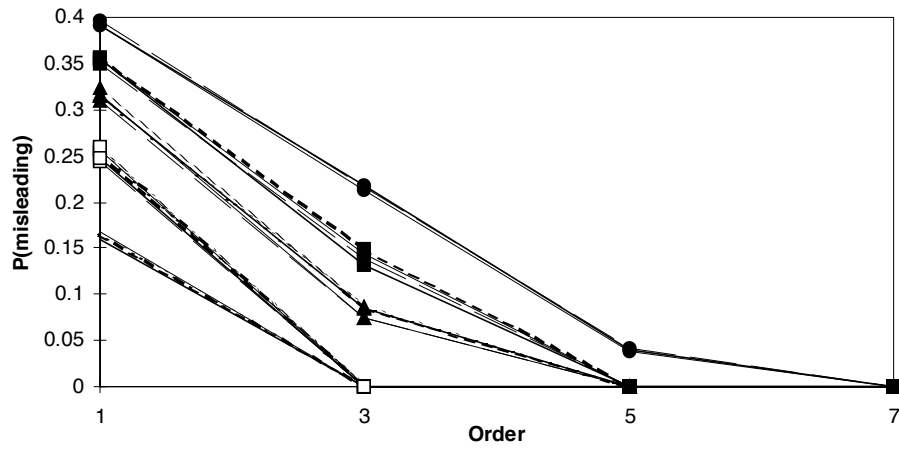


Figure 2: Probabilities of misleadingness, versus order, for landscapes where $N=8$. In this and the following figures, values of K are represented by markers: no marker = 2, white square = 3, triangle = 4, filled square = 5, circle = 6, diamond = 8. Values of P are represented by line type: light long dash = 2, light solid = 4, solid = 8, long dash = 16, short dash = 32, short-long dash = 64, short-short-long dash = 128 .

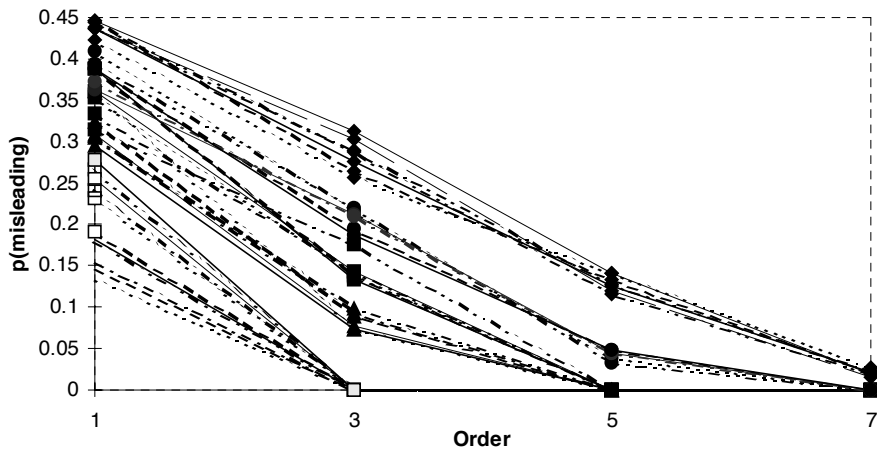


Figure 3: Probabilities of misleadingness, versus order, for landscapes where $N=16$. Symbols as noted in the previous figure.

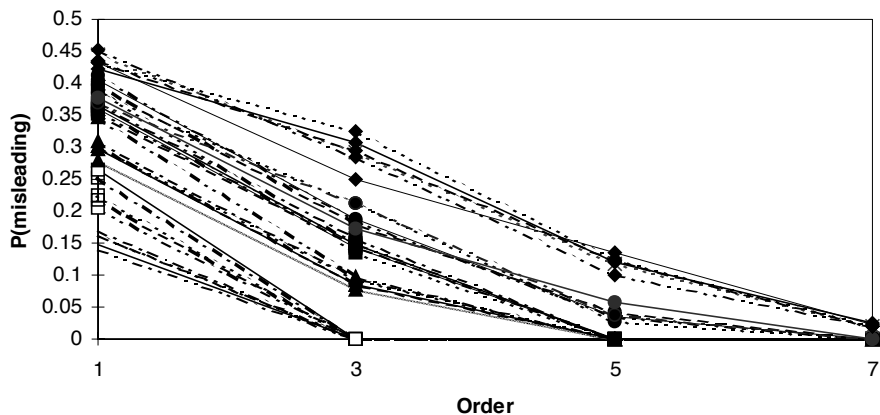


Figure 4: Probabilities of misleadingness, versus order, for landscapes where $N=20$. Symbols as noted in the previous figures.

Some general conclusions are worth noting. First, K bounds the level of misleadingness. For $n > K$, $P_m = 0$. For values of $n < K$, the probability increases with K , and decreases with n . Somewhat surprisingly, P seems to have no substantial effect on P_m . Note that the line groups for fixed N and K , and varying P , remain relatively the same. There are some variations, particularly for $N = 20$, but the absolute values of these variations are quite small. Finally, as was previously noted, the probabilities of misleadingness are bounded by 0.5. Therefore, the misleadingness is always less likely than a lack of misleadingness in the generated NKP landscapes. For higher orders of misleadingness, the graphs indicate that the probabilities are usually quite small.

10 Landscape Design

In addition to analysis, one can use the expressions presented here as a tool for landscape design. Consider using the Walsh-based procedure for constructing landscapes (outlined above), and manipulating the statistics of entries in the table **W**. As suggested by the previous analysis, the key parameters in such design are the Walsh sums, A and B . Parameter A represents the sum of odd Walsh coefficients of order greater than n , and parameter B represents the sum of odd Walsh coefficients of order less than n . Consider generating random variables A and B directly, from a bivariate normal distribution with known variances for A and B , and known correlation between A and B . Then, A and B could be broken down into Walsh coefficients, to fill table **W**.

It is important to note that a key parameter in such designs is ρ , the correlation between A and B . If negative correlation exists, the misleadingness probability is no longer bounded by 0.5. Therefore, the suggested Walsh-based procedure would allow for construction of landscapes with higher misleadingness probabilities. Note that although P seems to have little effect on juxtapositional complexity of problems created by the NKP procedure, it may have a more substantial effect when ρ is manipulated directly.

Finally, note that in the suggested procedure, the concepts represented by K (size of epistatic interactions), and P (number of epistatic interactions) remain. In this case, the size of epistatic interactions (K) is dictated by the largest order of non-zero Walsh coefficients, and the number of epistatic interactions (P) is dictated by the number of non-zero Walsh coefficients.

11 Final Comments

Several conclusions can be drawn from this study:

- In terms of peak distribution characteristics, the parameter P has little effect on the most critical of these characteristics.
- However, P acts as a “gain factor” on the landscape’s expected range of fitness values.
- Moreover, P seems more responsible than K for the so-called *complexity catastrophe* in these landscapes.
- The expected juxtapositional complexity (that is, the building block structure) of landscapes generated by the NKP procedure is limited.
- P seems to have little effect on changing the juxtapositional characteristics of landscapes generated by the NKP procedure.
- It may be possible to directly manipulate Walsh coefficient statistics to develop random landscapes with more desirable and controllable juxtapositional characteristics.

The last of these items will be the subject of further study. In particular, if such a procedure is followed to generate desired juxtapositional, the effects on peak distribution characteristics will have to be reconsidered.

As was discussed in the introduction, there are definite advantages to being able to generate and examine large numbers of random landscapes, if those landscapes can be parameterized in a meaningful way. Such landscapes can provide descriptive insight into real-world landscape complexity, and prescriptive insight as search algorithm test problems. It is hoped that this paper will aid in such efforts.

Acknowledgments

The authors thank Larry Bull for many useful discussions that contributed to this effort.

References

- Altenberg, L. (1994). Evolving better representations through selective genome growth. In *Proceedings of the 1st IEEE Conference on Evolutionary Computation*. pp. 182-187.
- Altenberg, L. 1997. NK fitness landscapes. In Bäck, T., Fogel, D., and Michalewicz, Z. (eds.) *The Handbook of Evolutionary Computation*. Oxford University Press. pp. B2.7:5-B2.7:10.
- Deb, K. 1997. Deceptive landscapes. In Bäck, T., Fogel, D., and Michalewicz, Z. (eds.) *The Handbook of Evolutionary Computation*. Oxford University Press. pp. B2.7:1-B2.7:5.
- Heckendorn, R. B., and Whitley, D. (1997). A Walsh analysis of NK-landscapes. In Bäck, T. (ed.) *Proceedings of the Seventh International Conference on Genetic Algorithms*. pp. 41-48. Morgan Kaufman.
- Hordijk, W. and Manderick, B. (1995). The usefulness of recombination. In Moran, F., Morelo, J. J., and Chacon, P. (eds.) *Advances in Artificial Life: Proceedings of the Third European Conference on Artificial Life*. Springer Verlag. pp. 908-919.
- Kauffman, S. A. (1993). *The Origins of Order: Self-Organization and Selection in Evolution*. New York: Oxford University Press.

Looney, S. W (1996). Sample size determination for correlation coefficient inference: Practical problems and practical solutions. Paper presented at 1996 Joint Statistical Meeting, August 4-8, Chicago, Illinois.

Maghsoodloo, S. and Huang, C. L. (unpublished). Computing probability integrals of a bivariate normal distribution. Unpublished technical report.

Smith, J. and Fogarty, T.C. (1996). Recombination strategy adaptation via evolution of gene linkage. In *Proceedings of the Third IEEE International Conference on Evolutionary Computing*. IEEE Press. pp. 826-831.

Weinberger, E.D. 1991. Local properties of Kauffman's N-K model, a tuneably rugged energy landscape. *Physical Review A*, 44(10). pp. 6399--6413.

