# A Walsh Analysis of NK-Landscapes

**Robert B. Heckendorn**
Department of Computer Science
Colorado State University
Fort Collins, Colorado 80523 USA
heckendo@cs.colostate.edu

**Darrell Whitley**
Department of Computer Science
Colorado State University
Fort Collins, Colorado 80523 USA
whitley@cs.colostate.edu

## Abstract

We use Walsh analysis to show NK-Landscapes form an extremely restricted set of functions out of the set of $N$ bit functions that have interactions between $K+1$ or fewer bits. Several upper bounds on the extent of the coverage by NK-Landscapes are created based on theory developed in the paper.

## 1 INTRODUCTION

A very useful and popular experimental model for correlated landscapes is Stuart Kauffman's NK-Landscapes [2, 3]. An NK-Landscape is a function $f : \mathcal{B}^N \rightarrow \mathcal{R}$ where $K$ is the number of bits in the chromosome that epistatically interact with each bit. One of the nice features of NK-Landscapes is that $K$ acts as a tunable ruggedness control. When $K = 0$, the landscape is the average of the sum of the weights associated with each bit and hence, assuming a Hamming neighborhood, is highly correlated and relatively smooth. When $K = N - 1$ the function is random and totally uncorrelated.

At first glance one might assume that an NK-Landscape can approximate any arbitrary function with up to $K + 1$ bits of interaction. However, NK-Landscapes were not designed for this purpose and in fact are much more restrictive. For example in a 20 bit function with up to 6 bits of interaction there can be $\binom{20}{6} = 38700$ different sources of 6 bit interactions. In an NK-Landscape with the same maximum level of interaction, $K = 5$, there can only be 20 such interactions. In fact, for many values of $N$ and $K$, NK-Landscapes represent an astronomically small subset of possible functions with $K + 1$ bits of interaction. We show that in many cases the coverage of Walsh coefficients is on the order of only $N$ in $2^{N-K}$. This dramatic difference prompts us to ask: how does the class of NK-Landscapes compare with the class of all possible $N$ bit functions with limited levels of interactions? We will use Walsh analysis to rigorously answer this question and to develop some intuition on how NK-Landscapes are constructed in general.

We begin by defining the concepts used in the paper and then present some theorems that will allow us to discuss an NK-Landscape by its image in Walsh coefficients. The reader that is interested in just the results may skip the rigor of the proofs by reading just the theorem statements and the surrounding text. In the next section we model an NK-Landscape and use that model to compare the landscape with the set of all possible functions of $N$ bits.

## 2 NOTATIONS AND DEFINITIONS

Let a **string** of length $L$ be an ordered list from the set $\mathcal{B} = \{0, 1\}$. The bits in a string are ordered $b_{L-1}, b_{L-2}, ...b_2, b_1, b_0$ and belong to $\mathcal{B}^L$. For example: `011001` is from $\mathcal{B}^6$ and has $b_0 = 1$. Nonnegative integers will be used interchangeably with binary for representing strings in $\mathcal{B}^L$ with $b_0$ being the least significant bit. A **hyperplane** or **schema** is denoted in the usual way. For example, hyperplane $h$ for strings in $\mathcal{B}^7$ might be `**1101*`.

- Let $[i]$ denote extracting the $i^{th}$ bit. So if $x$ =`1111011` then $x[2] = 0$. $x[i, j]$ denotes extracting and right justifying the bits $i$ through $j$. For example $x[2, 5]$ =`1110`.
- The **bit count** function $bc(i)$ returns the number of 1's in $i$. For example $bc(\texttt{001011}) = 3$ and $bc(15) = 4$.
- $i \subseteq j$ where $i, j \in \mathcal{B}^L$ reads as $i$ is **contained in** $j$. That is wherever there is a 1 in $i$ there is a 1 in $j$ or, said another way, $i \wedge \overline{j} = 0$.

- $\alpha$ and $\beta$ are defined on a hyperplane by the bit by bit replacements below as per Goldberg [1]:

$$\alpha(h) = \begin{cases} \{0,1\} & \rightarrow 1 \\ \{*\} & \rightarrow 0 \end{cases} \qquad \beta(h) = \begin{cases} \{1\} & \rightarrow 1 \\ \{*,0\} & \rightarrow 0 \end{cases}$$

For the hyperplane $h = **1101*$ :
$\alpha(h) = 0011110$ and $\beta(h) = 0011010$.

- $\langle f \rangle_h$ is the average of the values of the function $f$ over the domain of all strings in hyperplane $h$.

Walsh functions are used to analyze a function $f$, $f : \mathcal{B}^L \rightarrow \mathcal{R}$, that is the encoding of a problem domain in $L$ bit space as a real valued function.

$$f(x) = \sum_{i=0}^{2^L-1} w_i \psi_i(x)$$

is the **Walsh Polynomial** for $f$ where:

- $\psi_i(x) : \mathcal{B}^L \times \mathcal{B}^L \rightarrow \{-1, \ +1\}$ is the $i^{th}$ **Walsh Function** of $x$, as defined below.
- $w_i : \mathcal{B}^L \rightarrow \mathcal{R}$ is the $i^{th}$ **Walsh coefficient**. $w_i$ indicates the degree of interaction of bits indicated by the positions of the 1's in $i$. The **order** of $w_i$ is the number of one bits in $i$ (i.e. $bc(i)$).
- $w_i$ is defined to be the coefficient associated with the $i^{th}$ Walsh function $\psi_i(x)$. This concept will be used again and again throughout this paper.

This equation defines a mapping from the $2^L$ values in the table that defines the Walsh coefficients to another table of $2^L$ values that defines the function $f$. It is also important to note that the Walsh coefficients of the sum of two functions is the sum of the Walsh coefficients of the two functions.

A Walsh Function can be expressed as:

$$\psi_j(x) = Y(\bigoplus_{i=0}^{L-1}(x[i] \wedge j[i])) \qquad x, j \in \mathcal{B}^L$$

where $\bigoplus$ represents exclusive or and the function $Y$ performs the mapping: $0 \rightarrow 1, \ \ 1 \rightarrow -1$. From the symmetry in this expression it is obvious that $\psi_a(b) = \psi_b(a)$.

A zero value for a Walsh coefficient means that the bit combination represented by the index of the coefficient does not play a role in the computation of the function [5]. For example: if for all $i$ where the bit count $bc(i) = 5$, $w_i = 0$ then there are no 5 bit epistatic interactions. A measure of the diversity of interaction for a given number of interacting bits is the **Walsh count**, denoted by $\kappa_b$ where:

$$\kappa_b = \sum_{i:bc(i)=b} \text{NONZERO}(w_i)$$

with NONZERO returning 1 if $w_i$ is nonzero and 0 otherwise, and $bc(i)$ returning the bit count of $i$. $\kappa_b$ is, therefore, the number of nonzero Walsh coefficients for bit patterns $i$ with exactly $b$ bits set to 1. Notice that for an $L$ bit function there are $L+1$ Walsh counts. The vector $\vec{\kappa}$ can be thought of as the distribution of nonzero Walsh coefficients. The maximum value of $\kappa_b$ is $\binom{L}{b}$ and therefore for a totally random function the values in $\vec{\kappa}$ form a binomial distribution.

Let the **order of a function**, denoted $\Omega(f)$, be defined as the largest $i$ such that $\kappa_i \neq 0$. In the special case where all $\kappa_i$ are 0, that is $f(x) = 0$, $\Omega(f) = 0$. Intuitively, the order of a function is the size of the largest set of interdependent bits. So $\Omega(f)$ is a measure of the **maximum** number of bits of interaction found in $f$ and $\kappa_i$ measures the level of $i$-bit interaction.

# 3 THE EXPANSION THEOREMS

The Expansion Theorems will allow us to change the dimensionality of a function. We will use this in analyzing NK-Landscapes. But first we need to define two very useful operators: *pack* and *unpack*.

## 3.1 PACKING AND HYPERPLANE NUMBERING

A function *pack* is defined as $pack : \mathcal{B}^L \times \mathcal{B}^L \rightarrow \mathcal{B}^M$ where $M \leq L$. $pack(x, m)$ takes the bits in $x$ and masks them with a $L$ bit mask $m$ such that $bc(m) = M$ and packs the bits selected by the mask, right justified with zero fill, in the result. As we will see, *pack* can be used to map functions in $M$ bit space to functions in $L$ bit space.

A function *unpack* is defined as $unpack : \mathcal{B}^M \times \mathcal{B}^L \rightarrow \mathcal{B}^L$ where $M \leq L$ and an $L$ bit mask $m : bc(m) = M$. $unpack(x, m)$ takes all the bits in $x$ and unpacks them a bit in each position in the result that corresponds to a 1 in the mask. All unset bits are then set to 0. For example:

$$pack(10101, 01101) \Longrightarrow 011$$
$$unpack(011, 01101) \Longrightarrow 00101$$

Some observations about *pack* and *unpack* are:

- $unpack(pack(j, m), m) = j \wedge m$

- $pack(unpack(j,m),m) = j$
- $bc(pack(j,m)) = bc(j \wedge m)$
- $bc(unpack(j,m)) = bc(j)$
- $pack(i \wedge j, m) = pack(i,m) \wedge pack(j,m)$
- Given $m \in \mathcal{B}^L$, $j \in \mathcal{B}^M$, the set of $x \in \mathcal{B}^L$ such that $pack(x,m) = j$ is exactly the set of strings in $\mathcal{B}^L$ in an order $M$ hyperplane $h$ where $\alpha(h) = m$. The choice of $j$ and $m$ selects the particular order $M$ hyperplane.

With the last observation it becomes convenient to introduce a notation for numbering the hyperplanes. If $h$ is a hyperplane in $\mathcal{B}^L$ then we will use $h_{m,j}$, $j \in \mathcal{B}^M$, $m \in \mathcal{B}^L$, $bc(m) = M$, $M \leq L$ to denote a $M$ order hyperplane such that $x \in h_{m,j}$ if and only if $pack(x,m) = j$. This, in fact, is simply specifying the hyperplane by using $m$ to select the partition and $j$ to select the hyperplane from the partition.

For example: if $h_{m,j}$ specifies a hyperplane in $\mathcal{B}^7$ and if $j =$ `10110` and $m =$ `1101110` then $h_{m,j}$ is a $5^{th}$ order hyperplane contained in the partition `bb*bbb*`. $j$ specifies the fixed string positions. So $h(m,j)$ is `10*110*`. Some observations are:

- $\alpha(h_{m,j}) = m$
- $\beta(h_{m,j}) = unpack(j,m)$

### 3.2 FUNCTION EMBEDDING

We will show that NK-Landscapes can be viewed as a composition of $N$ lower dimensional functions. The next theorems are used to **embed** a lower dimensional function in a higher one using the *pack* and *unpack* functions.

Let $v(x) : \mathcal{B}^M \rightarrow \mathcal{R}$ and $f(x) : \mathcal{B}^L \rightarrow \mathcal{R}$ with $L \geq M$. $f$ is said to be an **expansion** of $v$ if there is a mask $m : bc(m) = M$ and $m \in \mathcal{B}^L$ such that $f(x) = v(pack(x,m)) \quad \forall x \in \mathcal{B}^L$. In short, expansion is how to embed a lower dimension function, $v$, in a higher dimension function, $f$, by setting the fitness of every string in the hyperplane $h_{m,x}$ of $f$ to $v(x)$. $m$ is referred to as the **expansion mask**. We denote $f$ as an expansion of $v$ using mask $m$ by:

$$f = \mathcal{E}(v,m)$$

The Walsh polynomials for two functions related by expansion are linear combinations of Walsh functions applied to two possibly different length pairs of strings. If $g$ is in $\mathcal{B}^M$ and $f$ is in $\mathcal{B}^L$ then the Walsh polynomial for $g$ has Walsh functions that take $M$ bit arguments while the Walsh polynomial for $f$ has Walsh functions that take $L$ bit arguments. The next two theorems show that the Walsh coefficients can be related by using *pack* and *unpack* functions.

**Theorem 1** (Pack/Unpack Equivalency)

For $i \in \mathcal{B}^M$ and $j \in \mathcal{B}^L$ and a mask $m \in \mathcal{B}^L$ with $bc(m) = M$:

$$\psi_j(unpack(i,m)) = \psi_{pack(j,m)}(i)$$

**Proof:**

$$\begin{aligned}
&\psi_j(unpack(i,m)) = \\
&\quad = Y(\textstyle\bigoplus_{n=0}^{L-1} unpack(i,m)[n] \wedge j[n]) \\
&\quad = Y(\textstyle\bigoplus_{n=0}^{L-1} (unpack(i,m) \wedge j)[n]) \\
&\qquad \text{pack with mask } m \text{ right justifies bits unpacked} \\
&\qquad \text{with mask } m \text{, so the parity in xor is the same:} \\
&\quad = Y(\textstyle\bigoplus_{n=0}^{M-1} pack(unpack(i,m) \wedge j), m)[n]) \\
&\quad = Y(\textstyle\bigoplus_{n=0}^{M-1} (pack(unpack(i,m),m) \wedge pack(j,m))[n]) \\
&\quad = Y(\textstyle\bigoplus_{n=0}^{M-1} (i \wedge pack(j,m))[n]) \\
&\quad = Y(\textstyle\bigoplus_{n=0}^{M-1} i[n] \wedge pack(j,m)[n]) \\
&\quad = \psi_{pack(j,m)}(i)
\end{aligned}$$

□

Note in the above theorem the left half is the Walsh function applied to $j$ and $unpack(i,m)$ which are both elements of $\mathcal{B}^L$, while on the right-hand side the arguments to the Walsh function are elements of $\mathcal{B}^M$.

Now we can apply this theorem in showing how function complexity as measured by $\Omega$ is affected by expanding a function.

**Theorem 2** (Packing Theorem)

Let $v(x) : \mathcal{B}^M \rightarrow \mathcal{R}$ and $f(x) : \mathcal{B}^L \rightarrow \mathcal{R}$ and $f = \mathcal{E}(v,m)$ then

$$w_i^f = \begin{cases} w^v_{pack(i,m)} & \text{if } i \wedge \overline{m} = 0 \\ 0 & \text{otherwise} \end{cases}$$

where $w^v$ and $w^f$ are Walsh coefficients for the $v$ and $f$ functions respectively and $i, m \in \mathcal{B}^L; bc(m) = M$.

**Proof:**

$$\begin{aligned}
w_i^f &= \tfrac{1}{2^L} \textstyle\sum_{x=0}^{2^L-1} f(x)\psi_x(i) \\
&= \tfrac{1}{2^L} \textstyle\sum_{x=0}^{2^L-1} v(pack(x,m))\psi_x(i) \\
&\quad \text{we change the dimension of the function here:} \\
&= \tfrac{1}{2^L} \textstyle\sum_{j=0}^{2^M-1} \sum_{x:pack(x,m)=j} v(pack(x,m))\psi_x(i) \\
&= \tfrac{1}{2^L} \textstyle\sum_{j=0}^{2^M-1} \sum_{x:pack(x,m)=j} v(j)\psi_x(i) \\
&= \tfrac{1}{2^L} \textstyle\sum_{j=0}^{2^M-1} v(j) \sum_{x:pack(x,m)=j} \psi_x(i) \\
&= \tfrac{1}{2^L} \textstyle\sum_{j=0}^{2^M-1} v(j) \sum_{x \in h_{m,j}} \psi_x(i) \\
&= \tfrac{1}{2^L} \textstyle\sum_{j=0}^{2^M-1} v(j) \sum_{x \in h_{m,j}} \psi_i(x)
\end{aligned}$$

We know from the Balanced Sum Theorem for Hyperplanes [1] that

$$\sum_{x \in h_{m,j}} \psi_i(x) = \begin{cases} 0 & \text{if } i \wedge \overline{\alpha(h_{m,j})} \neq 0 \\ \psi_i(\beta(h_{m,j}))|h_{m,j}| & \text{if } i \wedge \overline{\alpha(h_{m,j})} = 0 \end{cases}$$

where $|h_{m,j}|$ is the number of strings in $h_{m,j}$. But as we have seen $\alpha(h_{m,j}) = m$. So, starting where we left off and using the Balanced Sum Theorem:

CASE 1: Assume $i \wedge \overline{m} \neq 0$:

$$\begin{aligned} w_i^f &= \tfrac{1}{2^L} \textstyle\sum_{j=0}^{2^M-1} v(j) \sum_{x \in h_{m,j}} \psi_i(x) \\ &= \tfrac{1}{2^L} \textstyle\sum_{j=0}^{2^M-1} v(j) 0 \\ &= 0 \end{aligned}$$

CASE 2: Assume $i \wedge \overline{m} = 0$:

$$\begin{aligned} w_i^f &= \tfrac{1}{2^L} \textstyle\sum_{j=0}^{2^M-1} v(j) \sum_{x \in h_{m,j}} \psi_i(x) \\ &= \tfrac{1}{2^L} \textstyle\sum_{j=0}^{2^M-1} v(j) \psi_i(\beta(h_{m,j}))|h_{m,j}| \\ &= \tfrac{1}{2^L} \textstyle\sum_{j=0}^{2^M-1} v(j) \psi_i(\beta(h_{m,j})) 2^{L-M} \\ &= \tfrac{1}{2^M} \textstyle\sum_{j=0}^{2^M-1} v(j) \psi_i(\beta(h_{m,j})) \\ &= \tfrac{1}{2^M} \textstyle\sum_{j=0}^{2^M-1} v(j) \psi_i(unpack(j,m)) \\ &= \tfrac{1}{2^M} \textstyle\sum_{j=0}^{2^M-1} v(j) \psi_{pack(i,m)}(j) \\ &= w^v_{pack(i,m)} \end{aligned}$$

□

The Packing Theorem means that the only indices for the nonzero Walsh coefficients of an expansion of a function occur when all of the one bits for the index fall entirely in the expansion mask. Therefore, if $f = \mathcal{E}(v, m)$ then there are at most $\binom{bc(m)}{k}$ Walsh coefficients $w_i$ of $f$ where $w_i \neq 0$ and $bc(i) = k$. Specifically, for a mask $m$ such that $bc(m) = k$ there will be exactly one index for a nonzero Walsh coefficient of the expanded function with $k$ one bits set and it will be the one that exactly matches the expansion mask. This also means that all the nonzero Walsh coefficients lie within the hyperplane $h_{\overline{m},0}$ in the Walsh space.

**Corollary 1**

If $f = \mathcal{E}(v, m)$ then there is a 1-1 correspondence between $w^v_{pack(i,m)}$ and the potential nonzero Walsh coefficients of $f$, $w_i^f$.

**Proof:**

The proof is by showing there is a 1-1 correspondence between $i \in \mathcal{B}^L$ with $i \subseteq m$ and $j \in \mathcal{B}^M$ with $j = pack(i, m)$. It is clear from the Packing Theorem that for every $i$ such that $i \wedge \overline{m} = 0$ there exists exactly one $j$ that maps to it. Also the reverse function can be computed:

$$\begin{aligned} j &= pack(i, m) \\ unpack(j, m) &= unpack(pack(i, m), m) \\ unpack(j, m) &= i \wedge m \\ unpack(j, m) &= i \qquad \text{since } i \subseteq m \end{aligned}$$

Therefore each $j$ maps to exactly one $i$. Therefore the mapping is a 1-1 correspondence.

□

This allows us to show that even though an expansion of a function $f$ may be nonzero for every value in the domain, the $\Omega$ of the function is not increased. This concept is distilled in the following theorem.

**Theorem 3** (Expansion Theorem)

Let $v(x) : \mathcal{B}^M \rightarrow \mathcal{R}$ and $f(x) : \mathcal{B}^L \rightarrow \mathcal{R}$. If $f = \mathcal{E}(v, m)$ then $\Omega(f) = \Omega(v)$

**Proof:**

Since there is a 1-1 correspondence between the nonzero Walsh coefficients of $f$ and $v$; and since $bc(i) = bc(pack(i, m))$ for $i \subseteq m$, the sets of Walsh coefficients for any given number of interacting bits must be identical and hence $\Omega(f) = \Omega(v)$.

□

# 4 A WALSH ANALYSIS OF NK-LANDSCAPES

In this section we first analyze how NK-Landscapes cover the space of Walsh coefficients. We then show how to calculate the coverage more exactly if the masks are known and discuss the limitations of this computation.

## 4.1 AN UPPER BOUND ON WALSH COEFFICIENT COVERAGE

Using the notation we have developed, an NK-Landscape $f$ can be defined at a point $j$ as:

$$f_j = \frac{1}{N} \sum_{b=0}^{N-1} v_b(pack(j, m_b))$$

Each $v_b$ can be thought of as one of $N$ **interaction functions**, $v_b : \mathcal{B}^{K+1} \rightarrow \mathcal{R}$, that gives a partial fitness score for each bit pattern formed by the value of the $b^{th}$ bit itself and the $K$ bits it interacts with. Each bit may interact with a possibly **different** set of $K$ bits so $N$ interaction masks, $m_b$, are used to select the $K$ bits that epistatically interact with the $b^{th}$ bit. The $b^{th}$ bit itself is also selected by the mask. Therefore $bc(m_b) = K+1$. $K$ must therefore fall in the range $[0, N-1]$. The *pack* function uses the masks to select the state of the epistatic bits and generate the argument for the interaction functions $v_b$. Using this definition, $f_j$ can be thought of as an average of $N$ functions each of which is a composition of an interaction function and the *pack* function. This can be expressed in terms of an expansion as:

$$f_j = \frac{1}{N}\sum_{b=0}^{N-1} \mathcal{E}(v_b, m_b)$$

**Theorem 4** (NK-Landscape Limit Theorem)

If $f$ is an NK-Landscape then $\Omega(f) \leq K+1$.

**Proof:**

If $f$ is an NK-Landscape then it can be expressed as sum of expansion functions of $v_b$. Let $v'_b = \mathcal{E}(v_b, m_b)$, $v'_b : \mathcal{B}^L \rightarrow \mathcal{R}$. We know from the Expansion Theorem that the $\Omega(v'_b) = \Omega(v_b)$. From this it is clear that:

$$\Omega(f) \leq \max_{b=0}^{N-1} \Omega(v_b)$$

Since $v_b$ is a $K+1$ bit function it follows that $\Omega(v_b) \leq K+1$ and so $\Omega(f) \leq K+1$.

□

**Theorem 5**

An NK-Landscape never has more than $N$ nonzero Walsh coefficients of order $K+1$.

**Proof:**

Let $f$ be an NK-Landscape and $w_j^f$ be the Walsh coefficients of $f$.

$$f_j = \frac{1}{N}\sum_{b=0}^{N-1} v'_b(j)$$

where each $v'_b = \mathcal{E}(v_b, m_b), v'_b : \mathcal{B}^L \rightarrow \mathcal{R}$ and $\Omega(v'_b) \leq K+1$. Since adding functions is equivalent to adding the Walsh coefficients of those functions

$$w_j^f = \frac{1}{N}\sum_{b=0}^{N-1} w_j^b$$

where the $w_j^b$ are the Walsh coefficients of the $v'_b$. As a consequence of the Packing Theorem, since each $\Omega(v'_b) \leq K+1$ there is no more than one nonzero Walsh coefficient $w_j^b$ with $bc(j) = K+1$ for each function $v_b$. Hence there are no more than $N$ of them when the $v'_b$ functions are summed together.

□

This means that NK-Landscapes cannot necessarily generate all possible functions with an $\Omega$ of $K+1$, but only those with $N$ or fewer $K+1$ bit Walsh coefficients. This reasoning can be extended to show that the maximum number of Walsh coefficients of order $R$ in an NK-Landscape is limited by the minimum of the maximum number of possible order $R$ Walsh coefficients, $\binom{N}{R}$, and by the maximum contribution from the $N$ functions $N\binom{K+1}{R}$. This assumes no bit position has a 1 in more than one of the masks $m_b$. In practice there is a lot of duplication for $K > 0$ and the number of different Walsh coefficients of order $R$ is considerably less than our proposed upper bound of

$$\min(\binom{N}{R}, N\binom{K+1}{R})$$

The ratio of the maximum possible number of nonzero Walsh coefficients in an NK-Landscape, $N\binom{K+1}{R}$, to the total number of possible nonzero Walsh coefficients, $\binom{N}{R}$ at order $R$ is a good indicator of sparseness of the coverage by NK-Landscapes. Table 1 shows this comparison. The columns represent the varying values of $K$ for a 10 bit function. The rows represent the various orders, $R$, of Walsh coefficients for the function. Therefore each column represents a different NK-Landscape. Wherever a single number occurs in the table, the NK-Landscape may have the maximum number of nonzero Walsh coefficients for that order. When a ratio is given, the numerator is the upper bound on the number of nonzero Walsh coefficients for the NK-Landscape and the denominator is the maximum possible for any function of order $N$. For instance, for an NK-Landscape with $N = 10$ and $K = 5$ there are $\binom{10}{5} = 252$ Walsh coefficients of order 5 but at most $10\binom{6}{5} = 60$ of them can be nonzero in an NK-Landscape.

Another way to look at this is presented in figure 1. Here the X-axis represents the order, $R$, of the Walsh coefficients. The Y-axis represents the number of possible Walsh coefficients of that order. The standard Gaussian approximation for a binomial distribution is used to ease visualization.

The solid line represents the maximum number of possible nonzero Walsh coefficients for an arbitrary function over the 10 bit domain. This graph shows that the greatest number of possible Walsh coefficients are of order $N/2$ or 5. The upper dashed line is the maximum number of possible nonzero Walsh coefficients for

Table 1: The Upper Bound of Number of Nonzero Walsh Coefficients in a General 10 Bit Function vs the Number in an NK-Landscape ($N = 10$) Broken Down By Order R and $\Omega$ Limit $K + 1$.

| $\downarrow R \setminus K \rightarrow$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 2 | - | 10/45 | 30/45 | 45 | 45 | 45 | 45 | 45 | 45 | 45 |
| 3 | - | - | 10/120 | 40/120 | 100/120 | 120 | 120 | 120 | 120 | 120 |
| 4 | - | - | - | 10/210 | 50/210 | 150/210 | 210 | 210 | 210 | 210 |
| 5 | - | - | - | - | 10/252 | 60/252 | 210/252 | 252 | 252 | 252 |
| 6 | - | - | - | - | - | 10/210 | 70/210 | 210 | 210 | 210 |
| 7 | - | - | - | - | - | - | 10/120 | 80/120 | 120 | 120 |
| 8 | - | - | - | - | - | - | - | 10/45 | 45 | 45 |
| 9 | - | - | - | - | - | - | - | - | 10 | 10 |
| 10 | - | - | - | - | - | - | - | - | - | 1 |

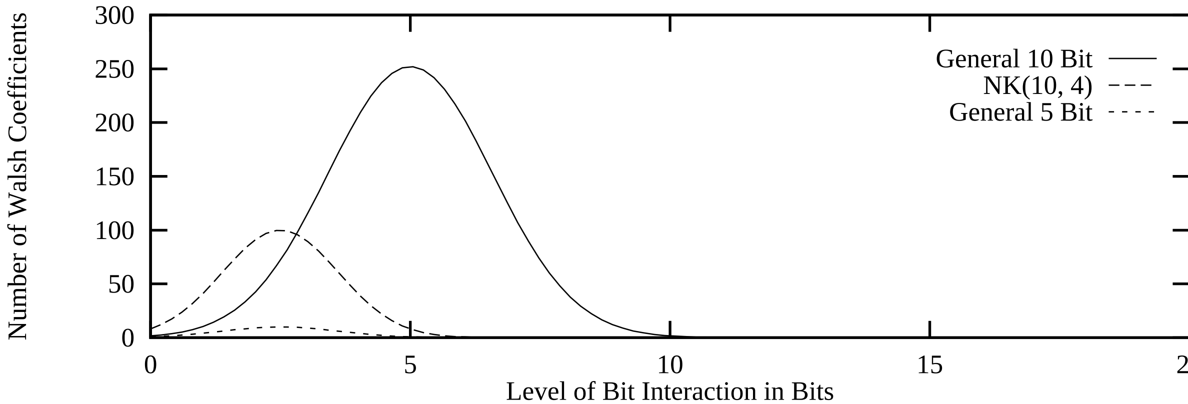


Figure 1: Walsh Coefficient Distributions for an Arbitrary 10 bit Function and for a (10, 4) NK-Landscape

an NK-Landscape with $N = 10, K = 4$. This function is the average of 10 functions whose maximum order of nonzero Walsh coefficients is 5. Each subfunction, denoted $v_b'$ above, has its own distribution of Walsh coefficients denoted by the lower dashed curve. When the 10 functions are averaged the coverage is increased 10 times and the left tail of the upper dashed curve arches above the solid curve which limits the total number of Walsh coefficients for the NK-Landscape. For larger $K$ the curve increases in size and the mean moves to the right until all possible functions of order $N-1$ are covered by a $K = N - 2$ order landscape. At $K = N - 1$ the curves no longer intersect and all possible functions can be generated. As $N$ increases the discrepancy between NK-Landscapes and arbitrary functions of $\Omega = K + 1$ becomes increasingly pronounced. In Figure 2 the solid curve is the potential number of nonzero Walsh coefficients for a 20 bit function while the small dashed curve at the bottom of the graph is the coverage for an NK-Landscape with $K = 9$.

In table 2 we give the intersection points and coverage for NK-Landscapes of $N = 40$ and $N = 80$ for various

Table 2: Coverage and Intersection Points for NK-Landscapes

| N=40 | | | N=80 | | |
|---|---|---|---|---|---|
| K | Coverage | I | K | Coverage | I |
| 0 | $1.00E{+}00$ | 2 | 0 | $1.00E{+}00$ | 2 |
| 4 | $1.42E{-}03$ | 2 | 8 | $1.52E{-}07$ | 2 |
| 8 | $5.21E{-}05$ | 3 | 16 | $7.65E{-}11$ | 3 |
| 12 | $1.53E{-}05$ | 4 | 24 | $4.22E{-}12$ | 4 |
| 16 | $2.21E{-}05$ | 5 | 32 | $7.81E{-}12$ | 5 |
| 20 | $1.11E{-}04$ | 6 | 40 | $2.31E{-}10$ | 7 |
| 24 | $1.27E{-}03$ | 8 | 48 | $3.79E{-}08$ | 9 |
| 28 | $1.93E{-}02$ | 10 | 56 | $9.54E{-}06$ | 12 |
| 32 | $2.76E{-}01$ | 16 | 64 | $2.44E{-}03$ | 19 |
| 36 | $9.98E{-}01$ | 28 | 72 | $5.27E{-}01$ | 36 |
| 39 | $1.00E{+}00$ | – | 79 | $1.00E{+}00$ | – |

values of $K$. The **coverage** is the ratio of the number of possible nonzero Walsh coefficients to available Walsh coefficients. The **intersection point**, $I$, is the intersection of the maximum possible coverage of $N$ independent $K+1$ order functions and the limit imposed by the general $N$ bit function. That is, intersection of upper dashed line and the solid line on the graph as

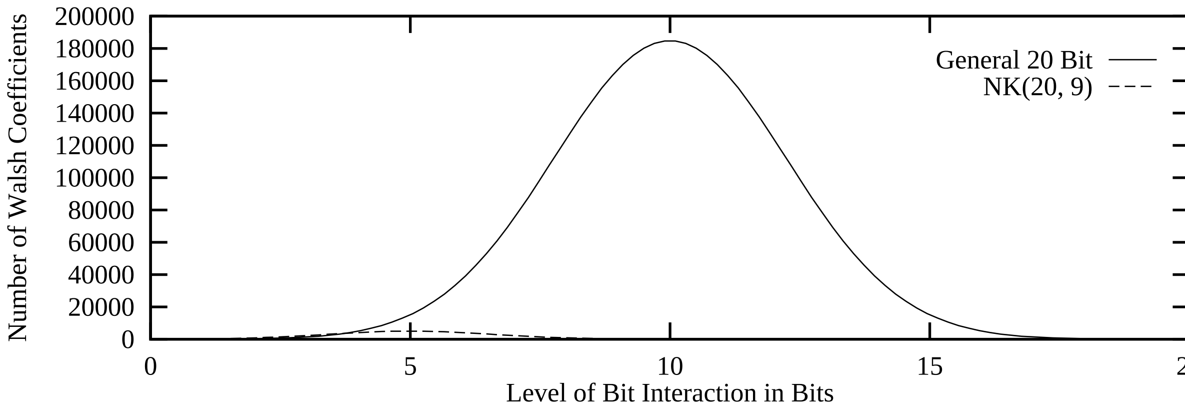


Figure 2: Walsh Coefficient Distributions for an Arbitrary 20 bit Function and for a (20, 9) NK-Landscape

measured in bits of interaction. Note that $I$ remains small for most values of $K$ and only nears $N$ when $K$ nears $N$. This allows the number of nonzero Walsh coefficients for an NK-Landscape to be approximated by $N2^{K+1}$ for most values of $K$ except near 0 and $N$. The number of available Walsh coefficients of a given order or less is approximately $2^{N-1}$ for $K \geq N/2$ therefore for $K \geq N/2$ and not near $N$ an upper bound for the coverage is $4N/2^{N-K}$. As K decreases from N/2 the number of coefficients for an NK-Landscape continues to decrease by powers of two while the number of available coefficients initially decreases more slowly. The result is that for most $K < N/2$ where K is not near 0 the value of the coverage is less than it is at $K = N/2$ giving an upper bound of $4N/2^{N/2}$.

## 4.2 COMPUTING THE COVERAGE OF WALSH COEFFICIENTS BY MASKS

In the previous section we showed an upper bound for the distribution of nonzero Walsh coefficients. This assumed that none of the $N$ bit positions has a 1 in more than one of the interaction masks $m_b$. If we know the set of masks we can better predict the upper bound of the distribution of Walsh coefficients. This gives us some insight on how overlapping regions of bit interaction interfere to reduce the diversity of functions.

We know from the Packing Theorem that the nonzero Walsh coefficients for the expansion of a function must have indices that are contained in the mask used for the expansion. That is, if $f = \mathcal{E}(v, m)$ then $w_i^f \neq 0$ only if $i \subseteq m$. Each interaction mask can therefore generate $2^{K+1}$ nonzero Walsh coefficients in the final function, but only if there is not any overlapping between masks.

But suppose there is overlapping. Let's examine the case of a function $f$ that is composed of the sum of the expansions of two functions, $f_1$ and $f_2$, based on corresponding overlapping masks $m_1$ and $m_2$ with $bc(m_1) = \delta_1$ and $bc(m_2) = \delta_2$. Assume the masks share $\delta_{12}$ bits in common, i.e. $bc(m_1 \wedge m_2) = \delta_{12}$. The maximum number of nonzero Walsh coefficients for function $f$ is the sum of the coefficients covered by the two functions minus the coefficients duplicated by the intersection.

Table 3: Example Walsh Count Computation

| $i$ | $\kappa_i^1$ | $\kappa_i^2$ | $\kappa_i^{12}$ | $\kappa_i^f$ |
|---|---|---|---|---|
| 0 | 1 | 1 | 1 | 1 |
| 1 | 4 | 4 | 2 | 6 |
| 2 | 6 | 6 | 1 | 11 |
| 3 | 4 | 4 | 0 | 4 |
| 4 | 1 | 1 | 0 | 1 |
| 5 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | 0 | 0 |

$$2^{\delta_1} + 2^{\delta_2} - 2^{\delta_{12}}$$

In fact because of the 1-1 correspondence established in the corollary to the Packing Theorem we know that the Walsh counts $\kappa_i$ for $f_1$ and $f_2$ for a given $i$ can only influence the order $i$ Walsh count for function $f$. Therefore the above upper bound can be stated more strongly as:

$$\kappa_i^f \leq \kappa_i^1 + \kappa_i^2 - \kappa_i^{12}$$

where $\kappa_i^1$ is the $i^{th}$ Walsh count for the function generated by an expansion based on $m_1$ and $\kappa_i^{12}$ is the Walsh count for a function generated by an expansion based on $m_1 \wedge m_2$.

For example: In table 3 we assume the function $f$ is composed of the expansion of two functions $f_1$ and $f_2$

on the domain $\mathcal{B}^4$ expanded with the respective masks $m_1 = \texttt{1110001}$ and $m_2 = \texttt{0011101}$. The first column is the order of the Walsh coefficients being counted. Since the masks are 7 bits long the domain of $f$ is $\mathcal{B}^7$ and therefore there are 8 different Walsh counts for $f$ and the expansions of $f_1$ and $f_2$. The second column is the distribution of the maximum number of nonzero Walsh coefficients for the expansion of $f_1$ given that the expansion was done using a mask with 4 bits set. $f_2$ is similarly represented in column three. The fourth column is the maximum number of nonzero walsh coefficients that are **duplicated** in the interaction of the two masks. The duplicate counts are subtracted and yield the final column which is the maximum number of nonzero coefficients for the given masks.

In general, by using the Inclusion-Exclusion Principle [4] from combinatorics we can see for $f$ based on N functions that are expanded using masks $m_b : b \in \{1, 2, ..., N\}$ the distribution of Walsh counts can be calculated as follows:

$$\vec{\kappa}^f = \sum_{1 \leq i \leq N} \vec{\kappa}^i - \sum_{1 \leq i < j \leq N} \vec{\kappa}^{ij} + \sum_{1 \leq i < j < k \leq N} \vec{\kappa}^{ijk} - \cdots - (-1)^N \vec{\kappa}^{(allmasks)}$$

The first term in the expression represents the sum of the contributions of each mask if the masks were all disjoint. For NK-Landscapes this only occurs with $K = 0$. Each successive sum in the expression is an interaction term for increasing numbers of masks.

Why doesn't this provide a formula for the exact number of nonzero Walsh coefficients? The answer is best illustrated by going back to the two mask example. The Walsh counts do not contain all of the information necessary to determine which Walsh coefficients are covered by both $f_1$ and $f_2$ when all possible Walsh coefficients of each order are **not** covered. Of course, an exhaustive listing of exactly which coefficients are covered by each expansion function and which are not, along with simple set operations, will determine the exact coverage.

To see how NK-Landscapes are assembled imagine an extension to the concept of NK-Landscapes called NKP-Landscapes that consist of $P$ masks, $m_b$, each $N$ bits long with $bc(m_b) = K+1$. With NKP-Landscapes we remove the restriction that $m_b$ has the $b^{th}$ bit set. We use this to illustrate what happens as the number of masks increases. When $P = 1$ the landscape consists of the expansion of a single subfunction $v$ by a mask $m$ resulting in every value in hyperplane $h_{m,x}$ being assigned $v(x)$. In Walsh space this makes the Walsh coefficients in the hyperplane $h_{\overline{m},0}$ the only nonzero Walsh coefficients. As $P$ increases the additive property of Walsh coefficients means that successive hyperplanes $h_{\overline{m_b},0}$ are added in the Walsh space while in function space, successive layers of constant hyperplanes $h_{m,x}$ are added.

If the restriction of having a fixed $K$ is removed, one can look at the reverse problem of expressing an $N$ bit function as a sum of functions of lower dimension. This can be viewed as decomposing a function into the sum of a series of hyperplanes in Walsh space. Our future research is directed to learning more about hyperplane competition using hyperplane decompositions in Walsh space.

## 5 CONCLUSION

We have shown that NK-Landscapes can be viewed as a synthesis of $N$ lower dimensional functions in function space and $N$ hyperplanes in Walsh space. Specifically, we showed that the expansion of a $K+1$ dimensional function produces a function with a predicatable signature in Walsh space determined by the expansion mask. We have also shown how the conflict between the interaction masks governs the coverage of the available Walsh coefficients and how truly sparse this coverage is. This paper does not say that NK-Landscapes are not useful. The user should be aware, on the one hand, that for most $K$, NK-Landscapes are a very small part of the possible functions of limited $\Omega$. On the other hand, this exact set of functions has shown a tremendous breadth of behavior as seen in Kauffman [2, 3] and that in itself is fascinating.

---

This research was supported by NSF grants IRI-9312748 and IRI-9503366.

## References

[1] David E. Goldberg, "Genetic Algorithms and Walsh Functions: Part I, A Gentle Introduction", *Complex Systems, Vol. 4, No. 3, 1989, pp129-152*

[2] Stuart A. Kauffman, "The Origins of Order", *Oxford Press, 1993*

[3] Stuart A. Kauffman, "At Home in the Universe", *Oxford Press, 1995*

[4] Ivan Niven "Mathematics of Choice", *Mathematical Association of America, 1965, pp67-77*

[5] Colin Reeves and Christine Wright, "An Experimental Design Perspective on Genetic Algorithms", *Foundations of Genetic Algorithms 3, 1995, pp7-22*