Finding the number of local optima in the 2-spin fitness landscape model

Angel Rodríguez

**Supervisor:** Peter Stadler

Bioinformatics Leipzig

*34th TBI Winterseminar in Bled*

February 12, 2019
Overview

1 Motivation.

2 Definition.

3 Properties.
   - Local Optima.

4 Random Field models for fitness landscapes.
   - Properties of Random Fields
   - Covariance matrix

5 Landscapes Models

6 Number of local optima
   - Local optima for the NK-model
   - Local optima for the 2-spin model
Discrete Problems

- Evolution and Genetic Algorithms.
- Physics of disordered systems.
- Discrete optimization problems.
A landscape consists of:

A set $V$ of configurations together with a notion of neighborhood:

And a ‘cost’ or ‘fitness’ function $f : V \rightarrow \mathbb{R}$

Therefore we can see $V$ as a graph $\Gamma$ with each vertex corresponding to one configuration and an edge between any two neighbors.
We will denote a configuration \( x \in V \) as \( x = (x_1, x_2, \ldots, x_n) \). We will call the entry \( x_i \) a **locus**, and it can only take the values +1 or -1. Therefore \( x \in \{\pm 1\}^n \).

**Neighborhood**

There is an edge between two configurations \( x, y \in V \) if they differ only in one entry. This makes our configuration space a Hamming graph \( H_2^n \).
Properties

Local and Global Optima
- The goal in landscapes problems is to find the ‘best’ value of the cost function $f$, which represent the fittest of individuals, the best tour in the TSP, or the lowest energy for spin glass. However local optima are obstacles in the way.
- A configuration $x \in V$ is a **local minimum** if $f(x) \leq f(y)$ $\forall y$ neighbor of $x$.
- And the minimum $x$ is called **global** if $f(x) \leq f(y)$ $\forall y \in V$.

Ruggedness
- The **number of local optima** is a measure for the ruggedness of a landscape. In general, it’s hard to calculate.
Random Field models for fitness landscapes

- We cannot hope to describe landscapes from their underlying biological and chemical structure, therefore one approach taken is to consider probabilistic models of fitness landscapes.

**Definition**

The set \( \{ f : V \rightarrow \mathbb{R} \} \) together with a measure \( \mu \{ f \} \) form the probability space \( \Xi \), which we call a **random field** on the graph \( \Gamma \). This measure can be seen as the form \( P(c_1, c_2, \ldots, c_{|V|}) \) which is the probability that for all configurations \( x_i \) holds simultaneously \( f(x_i) \leq c_i \), where \( c_i \in \mathbb{R} \).
Properties of Random Fields

**Expected values**

The expected value of a random variable $X$ defined on the random field is given by

$$E[X] := \int_{\mathbb{R}^n} X dP(c_1, \ldots, c_n)$$

**Covariance matrix**

The covariance matrix $C$ of a random field is defined componentwise as

$$C_{x,y} = \text{Cov}[f(x), f(y)] = E[f(x)f(y)] - E[f(x)]E[f(y)]$$

and is symmetric and non-negative definite.
Landscapes Models

House of Cards (HoC)
One of the simplest models. The fitness values \( f(x) \) for each configuration \( x \) are assigned independently at random from some probability distribution.

Sherrington Kirkpatrick (or 2-spin model)
This models a spin glass of \( n \) particles with 2 spin states, represented by \{+1, −1\}:

\[
H_{SK}(x) := \sum_{i<j} J_{ij} x_i x_j
\]

where the coupling constants \( J_{ij} \) are i.i.d. Gaussian random variables with mean 0 and variance 1.

P-spin model
A generalization of the Sherrington Kirkpatrick model in which each locus interacts with another \( p-1 \) locus.
In this model, the fitness \( f(x) \) of a configuration \( x \) is given by

\[
f(x) = \sum_{i=1}^{n} f_i(x_{b_{i,1}}, x_{b_{i,2}}, \ldots, x_{b_{i,K}})
\]

where \( f_i \) is an independent HoC landscape for the locus \( i \), depending on \( K \) loci and the indices \( b_{i,j} \) represent which \( K \) loci are being considered.
Note that the values for all the $b_i$ were left undefined, and exactly these values model the interaction scheme between loci. For instance:

**Adjacent neighborhood**

Each sub-landscape $f_i$ depends on the $i$-th locus and its $K - 1$ following neighbors. That is,

$$b_i = (x_i, x_{i+1}, \ldots, x_{i+K})$$

each element modulo $n$.

**Random neighborhood**

The neighborhood set $b_i$ contains $i$ and $K - 1$ other numbers, which are chosen at random from \{1, 2, \ldots, n\}.
Interaction between loci (NK model)

Block neighborhood

With $n$ an integer multiple of $K$, $n$ is divided into $n/K$ disjoint $K$-subsets and each block effectively behaves as an independent HoC landscape.

Adjacent neighborhood  Block neighborhood  Random neighborhood
How to find the number of local optima?

The idea is to find the probability $\pi_{max}$ that a randomly chosen configuration $x$ is a maximum, and the actual number of maxima $\#_{max}$ is just

$$\#_{max} = 2^n \pi_{max}$$

For the NK-model we will try to find $1/2 \leq \lambda_{model}^k \leq 1$ such that

$$\#_{max} = (2 \lambda_{model}^k)^n$$

where the factor 2 is introduced just as a reminder that the search space $(\mathbb{H}_2^n)$ increases as $2^n$. 

Angel Rodríguez (Bioinformatics Leipzig)
Calculating the local optima

The easiest example is the House of Cards model. Since $f(x)$ is random for every configuration, $\pi_{max}$ is just the probability that $x$ is the maximum among $n + 1$ random values, which is

$$\pi_{max} = \frac{1}{n + 1}$$

However, there is a general formalism for computing $\pi_{max}$ in the HoC and other models. Let us define the operator $\Delta_l : H^2 \leftarrow H^2$ for all $l \in \{1, 2, \ldots, n\}$ as

$$(\Delta_l x)_m = (1 - 2\delta_{lm})x_m$$

which changes the $l$-th entry of the configuration $x$. 
Calculating the local optima

Let \( h_0 \) and \( h_l \) be the fitness values for \( x \) and \( \Delta_l x \) respectively, i.e.,

\[
h_0 = f(x) \quad \text{and} \quad h_l = f(\Delta_l x).
\]

Then, \( x \) is a local maximum if \( h_0 > h_l \) or \( u_l \equiv h_0 - h_l > 0 \) for all \( 1 \leq l \leq n \). With vector notation, \( \mathbf{u} \equiv (u_1, u_2, \ldots, u_n) \), the **joint probability density** of the \( u_l \) is given by

\[
P(\mathbf{u}) = \int \prod_{l=0}^{n} dh_l p_f(h_l) \prod_{l=1}^{n} \delta(u_l - (h_0 - h_l)),
\]

and the characteristic function

\[
\Phi(\mathbf{q}) = \int dy \ p_f(y) \left( \prod_{l=1}^{n} \phi_f(-q_l) \right) \exp \left( iy \sum_{l=1}^{n} q_l \right)
\]

where \( \phi_f(-q_l) \) is the individual characteristic function of \( p_f(h) \).
Calculating the local optima

By performing the inverse Fourier transform of $\Phi(q)$ and then integrating over only positive values of $u_l$ (condition for $x$ to be maximum and represented by $\theta(u > 0)$), we obtain:

$$\pi_{max} = \prod_{l=1}^{n} \int_{0}^{\infty} du_l \ P(u) = \int \frac{DuDq}{(2\pi)^n} e^{-i\mathbf{q} \cdot \mathbf{u}} \theta(\mathbf{u} > 0) \ \Phi(q)$$

With these expressions one can recover the number of local optima for the HoC model and more importantly, one can get the number of maximum for some neighborhood models of the NK model.
The idea:

1. The NK-model with $K = 2$ and the adjacent neighborhood is the 2-spin model. Find the number of maxima for this case.

2. Calculate the covariance matrix of the 2-spin model.

3. Compare the covariance values with the number of maxima and try to find a relation between them.
Part 1. Local maxima for the NK model

In the NK model, the total fitness $F(x)$ of a configuration $x$ is a sum of individual HoC fitness values defined on each particular block or neighborhood.

Since a characteristic function is the natural object when dealing with sums of independent random variables, the approach will be based on the characteristic function of the NK blocks, and using the notation from previous slides, it has the form

$$\Phi(q) = \prod_{r=1}^{N} \Phi_r(q)$$

Where $\Phi_r(q)$ denotes the characteristic function of $u$ within the NK block $B_r$. 
Part 1. Local maxima for the NK model

With the help of an incidence matrix notation $b_{l,r}$ that indicates the presence (absence) of a locus $l$ in a neighborhood set $r$, i.e. $b_{l,r} = 1(0)$ if $l \in B_r(l \notin B_r)$; the characteristic function $\phi_r$ can be rewritten as:

$$\Phi_r(q) = \int dy_r \ p_f(y_r) \prod_{l=1}^{n} \left[ \phi_f(-q_l) e^{iy_rq_l} \right]^{b_{l,r}}$$

Then, once the full characteristic function has been derived, $\pi_{max}$ is readily calculated by inverse Fourier transform the same way as before, and for the Adjacent Neighborhood reads as:

$$\pi_{max}^{AN} = \int DyP(y) \frac{DuDq}{(2\pi)^n} e^{-i\mathbf{q} \cdot \mathbf{u}} \theta(\mathbf{u} > 0) \prod_{l=1}^{n} \phi_f(-q_l)^k e^{iq_l \sum_{r=0}^{k-1} y(l+r)_{\text{mod}(n)}}$$
TODO

- Simplify for $k = 2$, and compare with other results for 2-spin.
- Compare with covariance matrix.
Thank you.