# Linear Operators on Correlated Landscapes

By

Peter F. Stadler

Santa Fe Institute

1660 Old Pecos Trail, Santa Fe, NM 87501, USA

Institute for Theoretical Chemistry, University of Vienna*

Währingerstraße 17, A-1090 Vienna, Austria

Email: studla@itc.univie.ac.at or stadler@santafe.edu

*Address for Correspondence. Phone: **43 1 40480 678, Fax: ** 43 1 40 28 525

**PACS numbers:** 0250, 0550

**Abstract**

In this contribution we consider the effect of a class of "averaging operators" on isotropic fitness landscapes. Explicit expressions for the correlation function of the averaged landscapes are derived. A new class of tunably rugged landscapes, obtained by iterated smoothing of the random energy model, is established. The correlation structure of certain landscapes, among them the Sherrington-Kirkpatrick model, remains unchanged under the action of all averaging operators considered here.

## 1. Introduction

Evolutionary adaptation as well as combinatorial optimization take place on "landscapes" that result from mapping (micro)configurations to scalar quantities like fitness values or energies, or, more generally, to nonscalar objects like structures. The assumption of a particular mechanism for mutation, or the choice of a specific move-set for interconverting configurations, introduces the notions of neighborhood and distance between configurations. Viewing the set of all configurations, i.e., the configuration space, as a nondirected graph $\Gamma$ has turned out to be particularly useful. Each configuration is a vertex of $\Gamma$, and neighboring configurations are connected by edges. When configurations are sequences of equal length and the neighborhood relations are defined by the number of differing positions, one obtains the sequence space with the Hamming metric.

Among the most important properties of a landscape is its "ruggedness", which influences the speed of evolutionary adaptation as well as the quality of solutions of heuristic optimization strategies. While ruggedness has never been precisely defined, a number of empirical measures have been proposed, such as the number of local optima or the average length of up- or downhill walks [1,2], and the pair-correlation as a function of distance in configuration space [3,4].

The very notion of "uphill" or "optimum" implies a mapping to a scalar quantity. The definition of a pair correlation, however, can be extended to mappings to arbitrary objects, as long as there is a metric distance measure defined for them. Extensive studies on folding of RNA sequences into secondary structures have been based on this fact [5,6].

The landscapes of a number of well known combinatorial optimization problems, such as the Traveling Salesman Problem [7], the Graph Bipartitioning Problem [8], and the Graph Matching Problem [9]. Detailed information on the distribution of local optima and the statistical characteristics of downhill walks have been obtained for the uncorrelated landscape of the random energy model [10-12]. Furthermore, two one-parameter families of tunably rugged landscapes have been studied in detail: the $Nk$ model and its variants [2,5,13,14] and the $p$-spin models [15,16].

Here we will be concerned with the effect of averaging procedures related to running averages on landscapes. After some formal definitions (sect. 2) we discuss briefly the Fourier decomposition of landscapes (sect. 3) which is used in sect. 4 to derive the basic properties of linear operators applied to landscapes. In sect. 5 we briefly address the relation of averages over ensembles of landscapes and averages over configurations in single instances of landscapes. In sect. 6 particular operators, namely weighted averages over one-step neighborhoods, are used to construct a novel tunable family of landscapes. In sect. 7 the behavior of $p$-spin and $Nk$ landscapes under the iterated action of these operators is investigated.

## 2. Some Definitions

Let $\Gamma$ be a graph with $N < \infty$ vertices. A graph automorphism $\alpha$ is a one-to-one mapping of $\Gamma$ onto itself such that the vertices $\alpha(x)$ and $\alpha(y)$ are connected by an edge if and only if $x$ and $y$ are connected by an edge. Two vertices of $\Gamma$ are said to be equivalent if there is an automorphism $\alpha$ such that $y = \alpha(x)$. A graph is *vertex transitive* if all vertices are equivalent. A vertex transitive graph is always regular, i.e., each vertex has the same number of neighbors. The distance $d(x, y)$ of two vertices $x, y \in \Gamma$ is defined as the minimum number of edges that separate them. This distance is a metric. A graph is *distance transitive* if for any two pairs of vertices $(x, y)$ and $(u, v)$ with $d(x, y) = d(u, v)$ there is an automorphism $\alpha$ such that $\alpha(x) = u$ and $\alpha(y) = v$ (see, e.g., [17]). A graph $\Gamma$ is defined by its *adjacency matrix* $A$, with $A_{xy} = 1$ if the vertices $x$ and $y$ are connected by an edge, and $A_{xy} = 0$ otherwise.

In the following we will only consider graphs that are at least vertex transitive. Following Bollobás [18] we may then partition the vertex set by the following procedure: Choose an arbitrary vertex as reference point, which in the following will be labeled 0. Then construct pairwise disjoint sets of vertices $V(0) = \{0\}$, $V(1)$, $V(2)$, ..., $V(M)$ such that

$$\hat{a}_{\mu\nu} := \sum_{x \in V(\mu)} A_{xy} \qquad \text{is independent of } y \in V(\nu). \tag{1}$$

In other words, this partitioning is chosen such that each vertex $y$ in $V(\nu)$ is adjacent to $\hat{a}_{\mu\nu}$ vertices in $V(\mu)$. For some examples see figure 1 in section 3. It is induced by the symmetry of $\Gamma$: the vertices in each of the sets $V(\mu)$ are permutated by automorphisms that leave 0 invariant. In the worst case each set contains a single vertex. In general we will use greek letters to index the sets of such a partition, while roman indices are reserved for the vertices of $\Gamma$.

For highly symmetric graphs the *collapsed adjacency matrix* $\hat{A} = (\hat{a}_{\mu\nu})$ can be much smaller than the adjacency matrix $A$. In case of distance transitive graphs, for instance, the distance classes (i.e., the sets of all vertices with common distance from the reference vertex 0) form a partition fulfilling equ.(1), and the collapsed adjacency matrix $\hat{A}$ coincides with the intersection matrix as defined by Biggs [19] (see also [20]). We will use $d$ instead of greek letters for indexing the distance classes.

The following property of collapsed adjacency matrices will be used in the subsequent sections.

**Lemma 1.** $\sum_{y\in V(\kappa)} (A^m)_{xy} = (\hat{A}^m)_{\mu\kappa}$ for all $x \in V(\mu)$, and all non-negative integers $m$, and hence for each polynomial $p$ holds $\sum_{y\in V(\kappa)}[p(A)]_{xy} = [p(\hat{A})]_{\mu\kappa}$, i.e., $\hat{p}(A) = p(\hat{A})$.

**Proof.** We proceed inductively: The assertion is true by definition for $m = 1$. For $m = 0$ it holds trivially. Now suppose it is true for $m-1$; then

$$\sum_{y\in V(\kappa)} (A^{m-1}A)_{xy} = \sum_{\nu} \sum_{z\in V(\nu)} \sum_{y\in V(\kappa)} (A^{m-1})_{xz} A_{zy} =$$

$$= \sum_{\nu} \sum_{z\in V(\nu)} (A^{m-1})_{xz} \sum_{y\in V(\kappa)} A_{zy} = \sum_{\nu} (\hat{A}^{m-1})_{\mu\nu} \hat{A}_{\nu\kappa} = (\hat{A}^m)_{\mu\kappa}$$

The generalization to polynomials is straight forward. ∎

**Definition 2.** A *landscape* $f : \Gamma \to \mathbb{R}$ is a random field $\mathcal{F}$ on the set of vertices of $\Gamma$ defined by the distribution function

$$P(y_1, y_2, \ldots, y_N) = \text{Prob}\left\{f(x_i) \le y_i, \quad 1 \le i \le N\right\} \tag{2}$$

where $N$ is the number of vertices of $\Gamma$.

Let $\mathcal{E}[.]$ denote the mathematical expectation. For example, the expected value of the product of the values $f(x_k)$ and $f(x_l)$ at the vertices $x_k$ and $x_l$ is

$$\mathcal{E}[f(x_k)f(x_l)] = \int y_k y_l dP(y_1, y_2, \ldots, y_N). \tag{3}$$

**Definition 3.** A landscape is *isotropic* if

(i) $\mathcal{E}[f(x)] = \bar{f}$ for all configurations $x \in \Gamma$;

(ii) for any two pairs of configurations $(x, y)$ and $(u, v)$ such that there is a graph automorphism $\alpha$ with $\alpha(x) = u$ and $\alpha(y) = v$ holds $\mathcal{E}[f(x)f(y)] = \mathcal{E}[f(u)f(v)]$.

The second condition means that any two pairs of configurations that are equivalent in configuration space have the same pair-correlation. The covariance matrix of $\mathcal{F}$ is defined by

$$C_{xy} = \mathcal{E}[f(x)f(y)] - \mathcal{E}[f(x)]\mathcal{E}[f(y)]. \tag{4}$$

By vertex transitivity there is an automorphism $\alpha$ such that $\alpha(x) = 0$. Isotropy hence implies $C_{xy} = C_{0,\alpha(y)} = c(\mu)$, where $V(\mu)$ is the partition to which $\alpha(y)$ belongs. The variance is given by

$$\mathrm{Var}[f] = C_{xx} = C_{00} = c(0) \tag{5}$$

The *autocorrelation function* is defined as $\rho(\mu) = c(\mu)/c(0)$.

## 3. Fourier Series on Landscapes

Let $C$ be the covariance matrix of a random field $\mathcal{F}$ on some graph $\Gamma$. Denote by $\{\phi_s\}$ a complete set of orthonormal eigenvectors (which exists by symmetry of C). Although $\{\phi_s\}$ can be chosen real, we will admit complex vectors as well. Then the landscape $f$ on $\Gamma$ can be represented in mean square sense as

$$f(x) = \sum_{y\in\Gamma} a(y)\phi_y(x) \tag{6}$$

(The labeling of the eigenvectors $\phi_y$ with vertices is arbitrary.) This series representation is known as the *Karhunen-Loève* expansion.

**Remark.** The coefficients $a(y)$ are uncorrelated. For finite sets the Karhunen-Loève expansion coincides with the well known principal component analysis introduced by Hotelling [21] (see, e.g., [22]).

For an arbitrary graph $\Gamma$ with adjacency matrix $A$ the graph Laplacian is defined by $\Delta = D^{-1}A - E$, where $D$ is the diagonal matrix of vertex degrees (see, e.g., [20]). For vertex transitive graphs this becomes $\Delta = (1/D)A - E$, where $D$ is now the constant vertex degree of $\Gamma$. Denote by $\{\theta_s\}$ a complete orthonormal set of eigenvectors of $\Delta$. A series representation of the form

$$f(x) = \sum_{y\in\Gamma} b(y)\theta_y(x) \tag{7}$$

is called Fourier expansion of the landscape [2].

**Definition 4.** Let $(G, \circ)$ be a finite group, and let $\Phi$ be a set of generators of $G$ such that the group identity is not contained in $\Phi$, and for each $x \in \Phi$ the inverse group element $x^{-1}$ is also contained in $\Phi$. ($\Phi$ is a set of generators means that each group element $z \in G$ can be represented as a finite product of elements of $\Phi$, with multiplication defined by the group operation $\circ$.) Let $\Gamma(G, \Phi)$ be the graph with vertex set $G$ and an edge connecting two vertices $x$ and $y$ if and only if $xy^{-1} \in \Phi$, i.e., two vertices are connected if and only there is a $\gamma \in \Phi$ such that $y = \gamma x$. $\Gamma(G, \Phi)$ is called a *Cayley* graph (see, e.g., [23]).

**Remark.** A Cayley graph is vertex transitive.

A commutative group $(G, \circ)$ can be written as the direct product of a finite number $L$ of cyclic groups $C_j$ of order $N_j$. A group element can be represented by $x = (x_1, x_2, \ldots, x_L)$ with $0 \le x_j < N_j$; then the group operation becomes

$$z = x \circ y \quad\Longleftrightarrow\quad z_j = x_j + y_j \bmod N_j \qquad j = 1, \ldots, L. \tag{8}$$

It is easy to check that the functions $e_p$, $p \in G$, defined by

$$e_p(x) = \exp\left(2\pi i \sum_{j=1}^{L} \frac{x_j p_j}{N_j}\right), \tag{9}$$

fulfill $e_p(x)e_q(x) = e_{p\circ q}(x)$. Furthermore they are orthogonal with respect to the scalar product $\sum_{x\in G} e_q(x)e_q(x) = \delta_{pq}$. A simple consequence of these facts is the well known

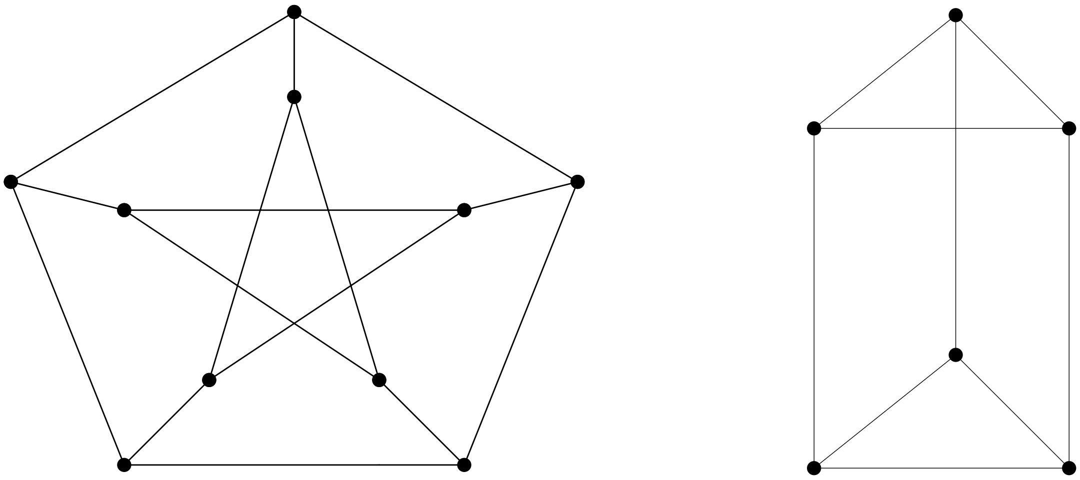
**Lemma 5.** Let $\Gamma(G,\Phi)$ be a Cayley graph of the commutative group $G$, and let $H_{ij} = h(j\circ i^{-1})$ depend only on the "difference" of the group elements $i$ and $j$. Then $e_p$ as defined in equ.(9) is an eigenvector of $H$. In particular, the set $\{e_p|p\in G\}$ forms an orthonormal basis of eigenvectors of the adjacency matrix of $\Gamma(G,\Phi)$.

**Remark.** Lovász [24] used this result to show that that the eigenvalues of the adjacency matrix of a Cayley graph $\Gamma(G,\Phi)$ with commutative group $G$ are given by $\lambda_p = \sum_{k\in\Phi} e_p(k)$.

Weinberger [14] showed using lemma 4 that Karhunen-Loève expansion and Fourier expansion coincide for isotropic landscapes on Cayley graphs arising from *commutative* groups. (He used a slightly more restrictive definition of isotropy, requiring the same covariance for all pairs of configurations with the same distance.) We will show that an analogous result holds for distance transitive configuration space.

**Remark.** There are graphs that are distance transitive but not Cayley graphs, and vice versa, there are Cayley graphs of commutative groups that are not distance transitive. The two most simple examples, the Petersen graph and the trigonal prism are shown in figure. 1.

**Lemma 6.** Let $a^{\ell}_{ij} = 1$ if $d(i,j) = \ell$, and 0 otherwise, for all $i,j\in\Gamma$. Denote the corresponding matrix by $A^{(\ell)}$. Let $\Gamma$ be distance transitive. Then $A^{(\ell)}$ is a linear combination of the first $\ell$ powers of $A$.



$$\hat{A} = \begin{pmatrix} 0 & 1 & 0 \\ 3 & 0 & 1 \\ 0 & 2 & 2 \end{pmatrix} \qquad \hat{A} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 2 & 1 \end{pmatrix}$$

**Figure 1:** Petersen's graph is distance transitive, but it is not a Cayley graph (see, e.g., [23]). It has three classes of vertices: the reference vertex (top), its neighbors, and all other vertices. Its collapsed adjacency matrix is given below the graph. The trigonal prism graph $T$ is the Cayley graph of $C_2{\times}C_3$ with set of generators $\Phi{=}\{(\gamma,e),(e,\varphi),(e,\varphi^{-1})\}$, with $C_2{=}\{e,\gamma\}$ and $C_3{=}\{e,\varphi,\varphi^{-1}\}$ and $e$ denoting the group identity. An edge connecting the two triangles is clearly not equivalent to an edge within a triangle, hence $T$ is not distance transitive. There are 4 symmetry classes of vertices: the reference vertex (top), its two neighbors in the upper triangle, its neighbor in the lower triangle, and the remaining two vertices, which have distance 2 from the reference.

**Proof.** The assertion is trivially true for $\ell = 0$ and $\ell = 1$, as $A^{(0)} = E$ and $A^{(1)} = A$. Now calculate $A \cdot A^{(\ell-1)}$. The triangle inequality assures that only pairs of indices $(i, j)$ with distances $\ell$, $\ell - 1$, and $\ell - 2$ can have non-zero entries. Distance transitivity of the graph guarantees that all entries belonging to the same distance class are equal, and hence

$$A \cdot A^{(\ell-1)} = c_+ A^{(\ell)} + c_0 A^{(\ell-1)} + c_- A^{(\ell-2)} \tag{10}$$

where $c_+ > 0$ as long as $\ell$ does not exceed the diameter of $\Gamma$. ∎

**Theorem 7.** Let $f$ be an isotropic landscape on a Cayley graph $\Gamma(G,\Phi)$ with a commutative group $G$, or on a distance transitive graph $\Gamma$. Then the adjacency matrix $A$ of $\Gamma$ and the covariance matrix $C$ of the landscape $f$ commute, $AC = CA$.

**Proof.** (a) For Cayley graphs of commutative groups the proposition is a direct consequence of lemma 5, see [14].

(b) If $f$ is isotropic and $\Gamma$ is distance transitive, then $C_{xy}$ depends only on the distance of the vertices $d(x,y)$. Hence $C$ can be written as a linear combination of the matrices $A^{(\ell)}$, and by lemma 6, it can be written as a polynomial in terms of $A$, $C = p(A)$.

In both cases $C$ and $A$ commute because they have the same eigenvectors. ∎

## 4. Linear Operators on Landscapes

Consider landscape $g$ obtained by some weighted averaging procedure from a landscape $f$, i.e., $g(x) = \sum_{y\in\Gamma} \Omega_{xy} f(y)$. In the following we will use the obvious matrix notations $g = \Omega f$.

**Lemma 8.** Let $C$ be the covariance matrix of the landscape $f$, and let $\Omega$ be a linear symmetric operator. Then the covariance matrix of $g = \Omega f$ is given by

$$C^{\Omega} = \Omega C \Omega \tag{11}$$

**Proof.** We assume without loosing generality that $\mathcal{E}[f_p] = 0$ for all $p \in \Gamma$.

$$\begin{aligned} C^{\Omega}_{xy} &= \mathcal{E}[(\Omega f)_x (\Omega f)_y] = \mathcal{E}\left[\sum_p \Omega_{xp} f_p \sum_q \Omega_{yq} f_q\right] = \\ &= \sum_{p,q} \Omega_{xp} \mathcal{E}[f_p f_q] \Omega_{qy} = (\Omega C \Omega)_{xy} \end{aligned}$$

∎

**Corollary 9.** Let $f$ be isotropic and assume that the configuration space $\Gamma$ is distance transitive or the Cayley graph of a commutative group. Suppose $\Omega$ is a polynomial of the adjacency matrix. Then we have $C^\Omega = (\Omega^2)C$.

**Proof.** From theorem 7 we know that $C$ and the adjacency matrix $A$ of $\Gamma$ commute. Thus $C$ commutes with all powers of $A$ and consequently with all polynomials of $A$. The corollary follows now immediately from lemma 8. ∎

**Theorem 10.** Let $f$ be isotropic with autocorrelation function $\rho$ and let $\Gamma$ be distance transitive (or a Cayley graph of a commutative group). Furthermore let $\Omega$ be some polynomial in $A$, $\Omega = P(A)$. Then $g = \Omega f$ is again isotropic on $\Gamma$ and

$$\rho^\Omega = c \cdot \hat{\Omega}^2 \rho \tag{12}$$

where the constant $c = 1/(\hat{\Omega}^2\rho)(0)$ is chosen such that $\rho^\Omega(0) = 1$.

**Proof.** From corollary 9 we know $C^\Omega_{0x} = (\Omega^2 C)_{0x}$. With $Q = \Omega^2$ we may write this as

$$C^\Omega_{0x} = C^\Omega_{x0} = \sum_\mu \sum_{y\in V(\mu)} Q_{xy} C_{y0} = \sum_\mu \hat{Q}_{\lambda\mu} C_{\mu 0}$$

for all $x \in V(\lambda)$, as a consequence of lemma 1. Dividing by the variance completes the proof. ∎

There exists a set of right eigenvectors of $\hat{A}$ that are orthogonal with respect to the weight function $w(\mu) = |V(\mu)|$ (for details see [25]). We may therefore expand the autocorrelation function with respect to this basis, $\rho(\mu) = \sum_\nu a_\nu \rho_\nu(\mu)$.

**Corollary 11.** Under the assumptions of theorem 10 we have

$$\rho^\Omega(\mu) = c \cdot \sum_\nu a_\nu P^2(\lambda_\nu)\rho_\nu(\mu) \tag{13}$$

where $\lambda_\nu$ denotes the eigenvalue of $\hat{A}$ belonging to $\rho_\nu$. In particular, if $\rho = \rho_j$ for some $j$, then we have $\rho^\Omega = \rho$.

## 5. Empirical Landscapes and Ensembles of Landscapes

The empirical variance $\sigma^2$ of an isotropic landscape $f$, i.e., of a particular instance of the random field $\mathcal{F}$ is measured by picking a sample of values at configurations randomly distributed on $\Gamma$. Denote by $p(\mu) = |V(\mu)|/N$ the probability that two randomly chosen vertices $x$ and $y$ belong to partition $\mu$, i.e., that there is an automorphism $\alpha$ with $\alpha(x) = 0$ and $\alpha(y) \in V(\mu)$. In case of a distance transitive configuration space we can interpret $p(d)$ as the probability to pick at random two configurations with distance $d$. The empirical variance can be written as [5]

$$\sigma^2 = \frac{1}{2N^2} \sum_{x,y} \left(f(x) - f(y)\right)^2 \tag{14}$$

and its expected value can be calculated:

$$\begin{aligned} \mathcal{E}[\sigma^2] &= \frac{1}{2N^2} \sum_{x,y} \left(2\mathcal{E}[f(x)^2] - 2\mathcal{E}[f(x)f(y)]\right) = \\ = \mathrm{Var}[f] - \sum_{\mu} p(\mu) \frac{1}{|V(\mu)|} &\sum_{y\in V(\mu)} \mathcal{E}[f(0)f(y)] = \mathrm{Var}[f]\left(1 - \sum_{\mu} p(\mu)\rho(\mu)\right) = \\ &= \mathrm{Var}[f](1-\bar{\rho}) \end{aligned} \tag{15}$$

Hence the expected empirical variance of a landscape $f$ is related to the ensemble variance, i.e., the variance of the random field $\mathcal{F}$ at a particular vertex, by

$$\mathcal{E}[\sigma^2] = \mathrm{Var}[f](1-\bar{\rho}) \tag{16}$$

The parameter $\bar{\rho}$ measures the average correlation of the landscape. As shown in [25] $\bar{\rho}$ is the contribution to $\rho$ corresponding to the largest eigenvalue of $A$, $\lambda_0 = D$.

The ensemble-autocorrelation function $\rho$ and the expectation of the empirical correlation function $\tilde{\rho}$ are hence related by

$$\tilde{\rho}(\mu) = \frac{\rho(\mu) - \bar{\rho}}{1-\bar{\rho}} \tag{17}$$

Equ.(17) implies that the omnipresent 'back-ground' correlation $\bar{\rho}$ has to be subtracted. We remark that for practical purposes it is the empirically observable correlation function $\tilde{\rho}(\mu)$ that matters. The autocorrelation function is said to be self-averaging if $\tilde{\rho} = \rho$ (see, e.g., [26]).

**Corollary 12.** Under the assumptions of theorem 10 it follows from $\rho = \tilde{\rho}$ that $\rho^{\Omega} = \tilde{\rho}^{\Omega}$, i.e., if the autocorrelation function of $f$ is self-averaging, then the autocorrelation function of the landscape $\Omega f$ is also self-averaging.

## 6. Iterated Smoothing Maps — A Tunable Family

By $N(x)$ we will denote the set of neighbors of $x$ plus $x$ itself, i.e., all vertices $y \in \Gamma$ with $d(x,y) \leq 1$. We define the *smoothing operator* $\Psi$ by

$$\Psi = \frac{1}{D+1}\,(A+E) \tag{18}$$

The smoothed landscape is then defined by $g = \Psi f$. Let $\Gamma$ be distance transitive. Then we obtain as an immediate consequence of theorem 10

$$\rho^{\Psi}(d) = \frac{\sum_{k=-2}^{2} \alpha_k(d)\rho(d+k)}{\sum_{k=0}^{2} \alpha_k(0)\rho(k)} \tag{19}$$

where $\alpha_k(d) = (\hat{\Psi}^2)_{d,d+k}$. The coefficients have a simple combinatorial interpretation: $\alpha_k(d)$ is the number of pairs $(u,v)$ with $u \in N(x)$, $v \in N(y)$ with $d(u,v) = d+k$ where $d(x,y) = d$ is fixed.

For particular graphs $\Gamma$ it is fairly easy to calculate the coefficients $\alpha_k(d)$ explicitly. Consider a general hypercube of diameter $n$ over an alphabet of size $\kappa$. The configuration space has $N = \kappa^n$ vertices. By distance transitivity we may choose

$$\begin{aligned} x &= (000...0000000....0000) \\ y &= (111...1111000....0000) \end{aligned} \tag{20}$$

with $d(x,y) = d$. We partition the set of pairs $(u,v)$ with $u \in N(x)$, $v \in N(y)$ into nine subsets depending on whether $u = x$, $u$ is neighbor of $x$ with the difference occurring in one of the first $d$ positions, or $u$ is a neighbor with the difference occurring in the latter $n-d$ positions, and the same for $v$. If $u=x$ and $v=y$ then $d(u,v)=d$. If $u$ differs form $x$ in one of the latter $n-d$ positions and $v=y$ then $d(u,v)=d+1$; there are $(n-d)(\kappa-1)$ such pairs. If $u$ differs from $x$ in the first section, however, there are $d(\kappa-2)$ pairs with distance $d$ and $d$ pairs with distance $d-1$, the latter arising from mutating a 0 to a 1. The combinations of mutations in both $x$ and $y$ can be treated similarly. Collecting all terms finally yields

$$\begin{aligned}
\alpha_{-2}(d) &= d(d-1) \\
\alpha_{-1}(d) &= 2d+(\kappa-2)(2d-1)d \\
\alpha_0(d) &= 1+(\kappa-1)(n+2d(n-d))+(\kappa-2)(\kappa d+d(d-1)(\kappa-2)) \\
\alpha_{+1}(d) &= 2(\kappa-1)(n-d)+(\kappa-2)(\kappa-1)(n-d)(2d+1) \\
\alpha_{+2}(d) &= (\kappa-1)^2(n-d)(n-d-1)
\end{aligned} \tag{21}$$

Let $f_0$ be the random energy model [27,28]. The autocorrelation function is $\rho(d)=\delta_{0,d}$ and $\tilde\rho(0)=1$, $\tilde\rho(d)=-1/(N-1)$ for $d>0$, respectively, where $N$ denotes the number of points in the configuration space. define $f_r=\Psi f_{r-1}$. Let $\tilde\rho^{(r)}$ denote the autocorrelation function of $f_r$. We will refer to this family of landscapes as *iterated smoothing landscapes*, ISLs.

Numerical evaluation of equation (19) with the initial condition $\tilde\rho$ on Boolean Hypercubes and their $\kappa$-letter generalizations shows that $\{f_r\}$ indeed forms a family of tunably rugged landscapes (fig. 2). In particular, we have

**Theorem 13.** Let $\Gamma$ be a generalized hypercube over an alphabet with $\kappa>2$ letters. Then

$$\lim_{r\to\infty}\tilde\rho^{(r)}(d)=1-\frac{1}{n}\frac{\kappa}{\kappa-1}d \tag{22}$$

On a Boolean Hypercube ($\kappa=2$) we have

$$\lim_{r\to\infty}\tilde\rho^{(r)}(d)=\frac{n}{n+1}(1-\frac{2d}{n})+\frac{1}{n+1}(-1)^d \tag{23}$$

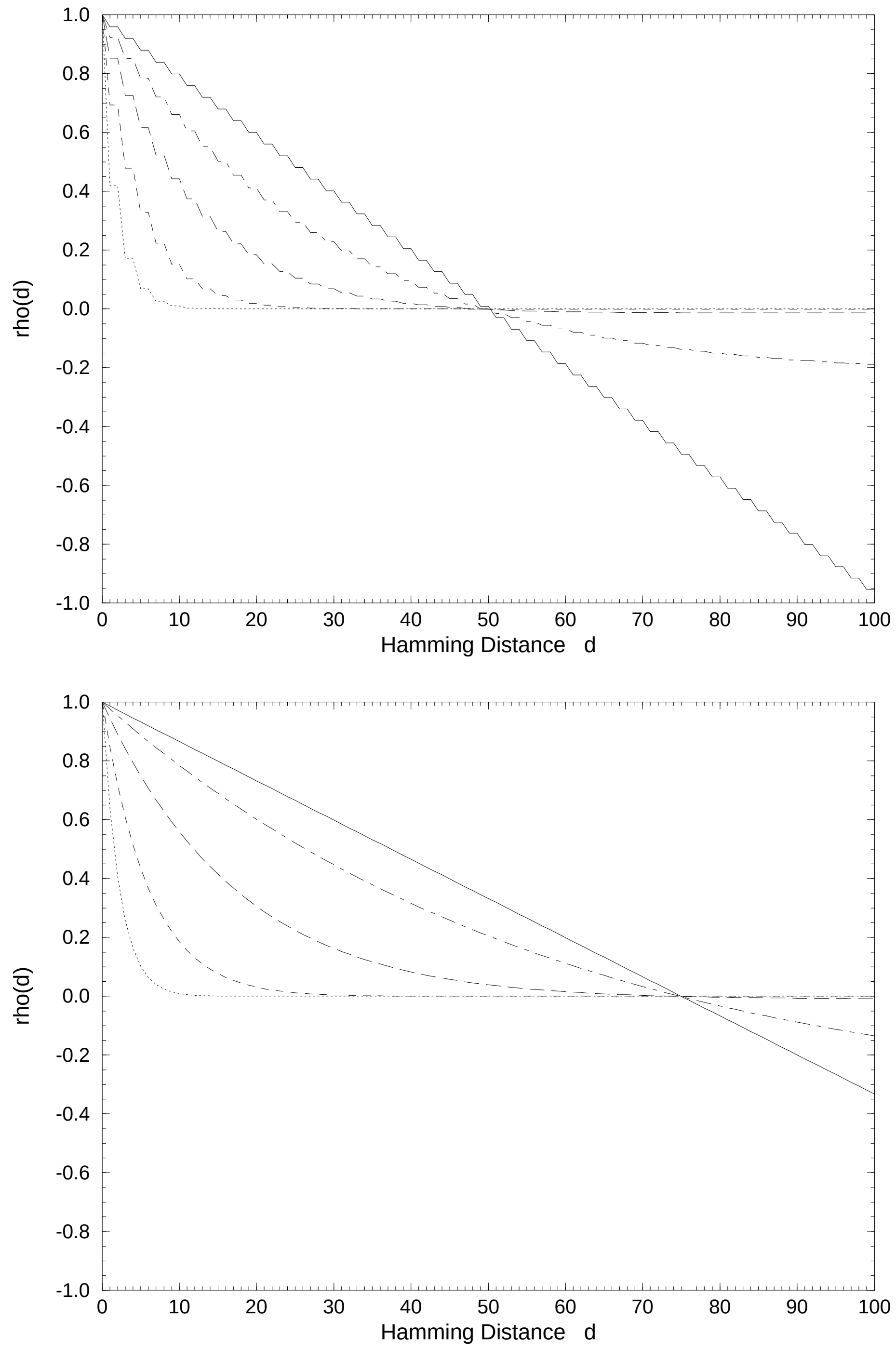



**Figure 2:** Autocorrelation function of ISLs on generalized hypercubes with **a)** $\kappa$=2, and **b)** $\kappa$=4 for with $n$=100. The number of smoothings are **a)** $r$=25 (dotted), $r$=50 (short dashed), $r$=75 (long dashed), $r$=100 (dot-dashed), $r$=200 (solid), and **b)** $r$=50 (dotted), $r$=100 (short dashed), $r$=150 (long dashed), $r$=200 (dot-dashed), and $r$=400 (solid), respectively.

**Proof.** The eigenvalues of the adjacency matrix of the general hypercubes are [29,30]

$$\lambda_j = n(\kappa - 1) - \kappa j \tag{24}$$

and hence the eigenvalues of the smoothing operator $\Psi$ are $1 - j\frac{\kappa}{n(\kappa-1)+1}$ for $j = 0, \ldots, n$. The largest eigenvalue of $\Psi^2$, apart from 1, which is obtained from $j = 0$, is found be setting $j = 1$ for $\kappa > 2$. In the case $\kappa = 2$ the second largest eigenvalue of $\Psi^2$ is degenerate, the corresponding eigenspace is spanned by $\rho_1$ and $\rho_n$. As an immediate consequence of equ.(28) in [25], the autocorrelation function of the random energy model can be written as

$$\tilde{\rho}^{(0)}(d) = c \cdot \sum_{\ell=1}^{n} (\kappa - 1)^{\ell} \binom{n}{\ell} \rho_{\ell}(d) \tag{25}$$

By corollary 11, the iterates of $\tilde{\rho}$ will converge to a vector in the eigenspace of the largest eigenvalue of $\hat{\Psi}^2$ that is not orthogonal to the initial condition. The eigenspace corresponding to $\lambda_0$ has been omitted by construction of $\tilde{\rho}$. By equ.(25) $\rho^{(0)}$, is not orthogonal to any of the eigenvectors $\rho_1$ through $\rho_n$. For $\kappa > 2$ hence, $\rho$ converges to $\rho_1$, while for $\kappa = 2$ we obtain a mixture of $\rho_1$ and $\rho_n$. The ratio of their contributions does not change under the action of $\Psi$ as a consequence of corollary 11. ■

We remark that the contribution of $\rho_n$ decreases with the size of the configuration space. The limiting autocorrelation function (22) corresponds to both the trivial spin glass model with Hamiltonian $\mathcal{H}(\sigma) = \sum_j A_j \sigma_j$, and to the $Nk$-model with $k = 0$. We also note that for $r/n \ll 1$ the landscapes are essentially uncorrelated, i.e., the correlation length is $o(n)$. Numerical estimates show that the nearest neighbor correlation $\tilde{\rho}^{(r)}(1)$ depends only on the ration $r/n$ for large $n$. (See figure 3.) As a consequence of theorem 13, we find that $\lim_{r\to\infty} \rho(1) = 1 - \frac{1}{n}\frac{\kappa}{\kappa-1}$ for $\kappa > 2$. For $\kappa = 2$ we find

$$\lim_{r\to\infty} \rho^{(r)}(d) = \lim_{r\to\infty} \rho^{(r)}(d-1) = 1 - \frac{2d}{n+1} \quad \text{for even } d \tag{26}$$

The data in fig. 3 indicate that this limit is already obtained for $r/n \geq 2$ to a good approximation.

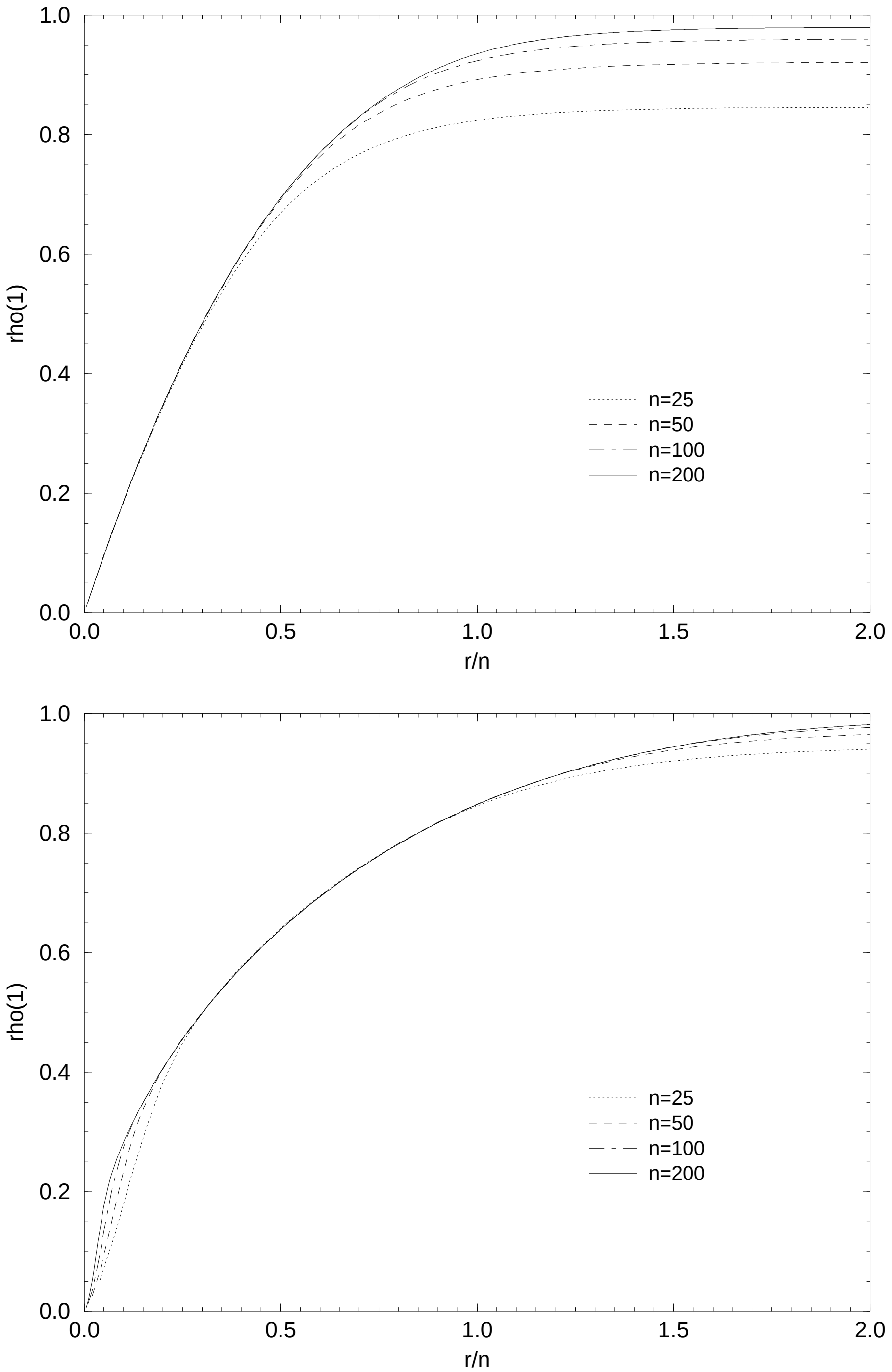



**Figure 3:** Nearest neighbor correlation for ISLs on generalized hypercubes as a function of the number of iterated smoothings. **a)** $\kappa=2$, **b)** $\kappa=4$.

## 7. Averaging Operators on Spin Glass and $Nk$ Landscapes

We first consider the $p$-spin models, i.e., the long range spin glasses with Hamiltonian

$$\mathcal{H}_p(\sigma) = \sum_{j_1<j_2<\ldots<j_p} A_{j_1 j_2 \ldots j_p}\, \sigma_{j_1}\sigma_{j_2}\ldots\sigma_{j_p} \tag{27}$$

with $\sigma_i \in \{-1,+1\}$ and coupling constants $A_{ij\ldots k}$ drawn from some common distribution. The ruggedness of the landscapes increases as the order $p$ of the spin-interactions increases. We have then

**Theorem 14.** Let $\Omega$ be any polynomial of the adjacency matrix of the Boolean Hypercube. Then the landscapes $\mathcal{H}_p$ and $\Omega\mathcal{H}_p$ have the same autocorrelation functions.

**Proof.** As shown in [16] the autocorrelation functions of the $p$-spin models are

$$\rho_p(d) = \binom{n}{p}^{-1} \sum_{\ell=0}^{n} (-1)^{\ell} \binom{d}{\ell}\binom{n-d}{p-\ell} \tag{28}$$

These are suitably normalized Krawtchouk polynomials, which are known to be the eigenvectors of the collapsed adjacency matrix of the Boolean Hypercube (see, e.g., [25,29]). Application of Corollary 11 completes the proof. ∎

In Kauffman's $Nk$-model the fitness of a binary string $x$ is defined as the average over "site fitnesses" $f_i$, which depend on $x_i$ and $k$ additional positions $i_1$ through $i_k$. The site fitnesses $f_i = f_i(x_i, x_{i_1}, \ldots, x_{i_k})$ are assumed to be uncorrelated random variables drawn from a common distribution [1]. The ruggedness of the landscape increases as the number $k$ of positions contributing to a single site fitness increases. Different variants of these landscapes have been investigated, depending on which positions contribute to a given site fitness (see, e.g., [31]).

In contrast to the $p$-spin models, the landscapes of $Nk$ type are affected by the smoothing operator $\Psi$.

**Theorem 15.** Let $f$ be an arbitrary $Nk$ landscape. Then the autocorrelation function of the landscapes $\Psi^r f$ converges to $\rho(d) = \xi(1-2d/n)+(1-\xi)(-1)^d$ with $0 < \xi \leq 1$ as $r$ tends towards infinity.

**Proof.** It is necessary and sufficient to show that the autocorrelation function $\rho$ of $f$ and $\rho_1$ are not orthogonal, i.e., that

$$\frac{1}{n}\sum_{d=0}^{n} \rho(d)\rho_1(d)\binom{n}{d} = \xi \neq 0 \tag{29}$$

Note that $\rho_1(d) = -\rho_1(n-d)$. Equ.(29) can hence be rewritten as

$$\xi = \frac{1}{n}\sum_{d<n/2} \binom{n}{d}\rho_1(d)\left[\rho(d) - \rho(n-d)\right]. \tag{30}$$

If $\rho(d) \geq \rho(n-d)$ for all $d < n/2$, and if the inequality holds strictly for at least one $d < n/2$, we have $\xi > 0$. It follows immediately from the very definition of the $Nk$ model (irrespective of how the neighboring sites are chosen) that $\rho$ is monotonically decreasing and non-constant. ∎

**Remark.** The variants of the $Nk$ model do not behave identically under the action of $\Psi$. Whenever the autocorrelation function is a polynomial of degree less than $n$, then we have $\xi = 1$ in theorem 15. For exact expressions of the autocorrelation functions of various $Nk$ models see [2,5]. Recently $Nk$-models on sequences of arbitrary alphabets have been considered [2]. Theorem 15 holds with $\xi = 1$ for all alphabets with three or more letters.

As a final example consider an isotropic landscape on the circle of length $n > 4$ (Millonas, pers. communication). The configuration space hence is the Cayley graph of the cyclic group $C_n = \{\gamma^k | 0 \leq k < n\}$ with set of generators $\Phi = \{\gamma, \gamma^{-1}\}$. Its vertex degree is $D = 2$. The eigenvectors of the adjacency matrix are $e_k(x) = \exp(\frac{2\pi k x}{n})$, as an immediate consequence of lemma 5.

The eigenvectors of the collapsed adjacency matrix are $\rho_k(d) = \cos(\frac{2\pi kd}{n})$ for $0 \leq k \leq n/2$, associated with the eigenvalues $\lambda_k = 2\cos(\frac{2\pi k}{n})$. The eigenvalue of $\hat{\Psi}^2$ corresponding to $\lambda_{n/2}$ is much smaller that the one corresponding to $\lambda_1$ for $n > 4$, hence $\rho$ converges to $\rho_1$ under the iterated action of $\Psi$ whenever $\rho$ is not orthogonal to $\rho_1$.

Suppose $f$ has a monotone non-increasing autocorrelation function $\rho$. As a particular example consider the random energy model on the circle. By the same reasoning as in the proof of theorem 15, $\rho$ is not orthogonal to $\rho_1$, and hence iterated smoothing eventually leads to a landscape with autocorrelation function $\rho_1$. Note that the smoothed landscapes on the Boolean Hypercube (and on related graphs) eventually have autocorrelation functions of the form $\rho(d) = 1 - \frac{1}{\ell}d + \mathcal{O}(d^2)$ while for the circle graphs we find $\rho(d) = 1 - \alpha d^2 + \mathcal{O}(d^3)$. The landscape on the circle hence becomes exceptionally smooth [16] as a consequence of the special geometry of the configuration space.

This example also shows that the degeneracy of $\Psi$ on the Boolean Hypercube is not a necessary consequence of the fact that the configuration space is bipartite, i.e., that there is an eigenvalue $-D$ of the adjacency matrix, since the circle graphs are bipartite for even $n$, and $\Psi$ is not degenerate there.

## 8. Discussion

A general theory of the action of linear (averaging) operators on isotropic landscapes is developed which allows to predict their effect on the correlation structure of the landscape. We have shown that there are particular landscapes (characterized by autocorrelation functions that are eigenvectors of the collapsed adjacency matrix of the underlying configuration space) which are not affected by "smoothing" operators. The $p$-spin models, and in particular the Sherrington-Kirkpatrick [32], model are members of this class on the Boolean hypercube. Another example is the landscape of the graph-bipartitioning problem.

A new family of tunably rugged landscapes is established, which arises by iterated smoothing of the random landscape. These "iterated smoothing landscapes" (ISLs) have an advantage over the other two known tunably rugged families: The $p$-spin models are defined only for Boolean hypercubes, and Kauffman's Nk family lives on general sequence spaces, while ISLs can be constructed on arbitrary configuration spaces. Their simple structure suggests furthermore that analytical solutions of these models might be feasible. Iterated smoothing of the random energy model eventually leads to a landscape with linearly decreasing autocorrelation function on general sequence spaces with alphabets of size $\kappa > 2$, while on Boolean hypercubes, $\kappa = 2$, one obtains a superposition of a linear function and an oszillation of the form $(-1)^d$. On a circle graph the same procedure leads to an autocorrelation function of the form $1 - \alpha d^2 + \ldots$. This shows again that the structure of a landscape cannot be separated from the geometry of the underlying configuration space.

Evolutionary adaptation can be described in the simplest case as the motion of a population in configuration space. Almost all analytical work, however, has been done with adaptive walks (or related processes) which neglect the population aspect. For a recent review see, e.g., [31]. To a first approximation a population can be replaced by its "center of mass" (consensus configuration), which moves in the landscape "experienced" by the entire population: again to a first approximation, this landscape is a weighted average of the fitnesses of configurations centered around the consensus. Consequently, the width of the population should strongly influence the dynamics of evolution: broader populations "feel" in general smoother landscapes, and hence optimization should easier. Usually the broadness of the population is controlled via the mutation rate [3,33] which at the same time controls the exploration rate. Further investigations are necessary in order to separate the effects of enhanced exploration and implicit smoothing of the underlying landscape caused by an increased mutation rate.

It should be kept in mind that the results presented here depend on the assumption of isotropic landscapes. While this requirement is often fulfilled for simple statistical models, it is unlikely to hold for real biophysical or physical systems. It has been shown, for instance, that the landscapes of RNA free energies are not isotropic for natural sequences [34].

Much work on random landscapes such as Derrida's REM has been motivated with the notion that a sufficiently "coarse grained" version of any rugged landscape would look random. More precisely, this is to say that it should be possible to construct renormalization groups having random landscapes as attractive fixed points. While the averaging procedures discussed in the this contribution are not exactly the kind of coarse graining one would have in mind for such a construction, the methods developed here seem to to be useful for such a task.

The theory outlined in this contribution is not applicable for the traveling salesman problem since the natural configuration spaces of the TSP are Cayley graphs of the symmetric group with the set of all transposition, or the set of all inversions (2opt-moves), as generators [7]. It is easy to check that these graphs are not distance transitive, and of course the symmetric group is non-commutative.

## Acknowledgements

I thank the Max Planck Society, Germany, for supporting my stay at the Santa Fe Institute. Stimulating discussions with P.W. Anderson, W. Fontana, S.A. Kauffman, M. Millonas, R. Palmer, and M. Virasoro are gratefully acknowledged.

## References


[1] Kauffman S.A. and S. Levin. Towards a General Theory of Adaptive Walks on Rugged Landscapes. *J. Theor. Biol.* **128**, 11-45 (1987).

[2] Weinberger E.D. Local Properties of the *Nk*-Model, a Tunably Rugged Landscape. *Phys. Rev. A* **44**, 6399-6813 (1991a).

[3] Eigen M., J.S. McCaskill, and P. Schuster. The Molecular Quasispecies. *Adv. Phys. Chem.* **75**, 149–263 (1989).

[4] Weinberger E.D. Correlated and Uncorrelated Fitness Landscapes and How to Tell the Difference. *Biol. Cybern.* **63**, 325-336 (1990).

[5] Fontana W, P.F. Stadler, E.G. Bornberg-Bauer, T. Griesmacher, I.L. Hofacker, M. Tacker, P. Tarazona, E.D. Weinberger, and P. Schuster. RNA Folding and Combinatory Maps *Phys. Rev. E* **47**, 2083–2099 (1993).

[6] Fontana W., D.A.M. Konings, P.F. Stadler, P. Schuster. Statistics of RNA Secondary Structures. *Biopolymers* **33**, 1389-1404 (1993).

[7] Stadler P.F. and W. Schnabl. The Landscape of the Traveling Salesman Problem. *Phys. Lett. A* **161**, 337-344 (1992).

[8] Stadler P.F. and R. Happel. Correlation Structure of the Graph Bipartitioning Problem. *J. Phys. A: Math. Gen.* **25**, 3103-3110 (1992).

[9] Stadler P.F. Correlation in Landscapes of Combinatorial Optimization Problems *Europhys. Lett.* **20** 479-482 (1992).

[10] Macken C.A., P.S. Hagan, and A.S. Perelson. Evolutionary Walks on Rugged Landscapes. *SIAM J. Appl. Math.* **51**, 799-827 (1991).

[11] Flyvbjerg H. and B. Lautrup. Evolution in a rugged fitness landscape. *Phys. Rev. A[15]* **46**, 6714-6723 (1992).

[12] Bak P., H. Flyvbjerg, and B. Lautrup. Coevolution in a rugged fitness landscape. *Phys. Rev. A[15]* **46**, 6724-6730 (1992).

[13] Kauffman S.A., E.D. Weinberger, and S.A. Perelson. Maturation of the Immune Response Via Adaptive Walks on Affinity Landscapes. In: *Theoretical Immunology, Part I*, ed.: Alan Perelson. Addison-Wesley, Redwood City, 1983.

[14] Weinberger E.D. Fourier and Taylor Series on Landscapes *Biol. Cybern.* **65**, 321-330 (1991b).

[15] Amitrano C., L. Peliti, and M. Saber. Population Dynamics in Spin Glass Model of Chemical Evolution. *J. Mol. Evol.* **29**, 513-525 (1989).

[16] Weinberger E.D. and P.F. Stadler. Why Some Fitness Landscapes are Fractal. *J. Theor. Biol.* **163**, 255-275 (1993).

[17] Buckley F. and F. Harary. *Distance in Graphs.* Addison-Wesley, Reading, MA, 1990.

[18] Bollobás B. *Graph Theory – An Introductory Course.* Springer-Verlag, New York, 1979.

[19] Biggs N. *Algebraic Graph Theory.* Cambridge University Press, Cambridge, 1974.

[20] Cvetković D.M., M. Doob, and H. Sachs. *Spectra of Graphs – Theory and Applications*, Academic Press, New York, 1980.

[21] Hotelling H. Analysis of a complex of statistical variables into principal components. *J. Educ. Psych.* **24**, 417-441 and 498-520 (1933).

[22] Rao C.R. *Linear Statistical Interference and Its Applications.* 2nd ed., Wiley, New York (1973).

[23] Biggs N.L. and A.T. White *Permutation Groups and Combinatorial Structures.* Cambridge Univ. Press, Cambridge, 1979.

[24] Lovász L. Spectra of graphs with transitive groups. *Periodica Math. Hung.* **6**, 191-195 (1975).

[25] Stadler P.F. Random Walks and Orthogonal Functions Associated with Highly Symmetric Graphs. *SFI preprint 93-06-31* (1993), *Disc. Math.* in press (1994).

[26] Binder K. and A.P. Young. Spin Glasses: Experimental Facts, Theoretical Concepts, and Open Questions. *Rev. Mod. Phys.* **58**, 801-976 (1986).

[27] Derrida B. Random Energy Model: An exactly solvable model of disordered systems. *Phys. Rev. B* **24**, 2613-2626 (1981).

[28] Gardner E. and B. Derrida. The probability distribution of the partition function of the random energy model. *J. Phys. A* **22**, 1975-1982 (1989).

[29] Dunkl C.F. Orthogonal Functions on Some Permutation Groups. In: Ray-Chaudhuri D.K. (ed.) *Proceeding of Symposis in Pure Mathematic* **34**. American Mathematical Society, New York 1979.

[30] Rumschitzky D. Spectral Properties of Eigen's evolution matrices. *J. Math. Biol.* **24**, 667-680 (1987).

[31] Kauffman S.A. *The Origins of Order.* Oxford University Press, New York, Oxford, 1993.

[32] Sherrington D., and S. Kirkpatrick. Solvable Model of a Spin Glass. *Phys. Rev. Lett.* **35**, 402-416 (1975).

[33] Eigen M. Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften* **58**, 465-523 (1971)

[34] Stadler P.F. and W. Grüner. Anisotropy in Fitness Landscapes. *J. Theor. Biol.* **165**, 373-388 (1993).

## Figure Captions

**Figure 1.** Petersen's graph is distance transitive, but it is not a Cayley graph (see, e.g., [23]). It has three classes of vertices: the reference vertex (top), its neighbors, and all other vertices. Its collapsed adjacency matrix is given below the graph.

The trigonal prism graph $T$ is the Cayley graph of $C_2 \times C_3$ with set of generators $\Phi = \{(\gamma, e), (e, \varphi), (e, \varphi^{-1})\}$, with $C_2 = \{e, \gamma\}$ and $C_3 = \{e, \varphi, \varphi^{-1}\}$ and $e$ denoting the group identity. An edge connecting the two triangles is clearly not equivalent to an edge within a triangle, hence $T$ is not distance transitive. There are 4 symmetry classes of vertices: the reference vertex (top), its two neighbors in the upper triangle, its neighbor in the lower triangle, and the remaining two vertices, which have distance 2 from the reference.

**Figure 2.** Autocorrelation function of ISLs on generalized hypercubes with **a)** $\kappa = 2$, and **b)** $\kappa = 4$ for with $n = 100$. The number of smoothings are **a)** $r = 25$ (dotted), $r = 50$ (short dashed), $r = 75$ (long dashed), $r = 100$ (dot-dashed), $r = 200$ (solid), and **b)** $r = 50$ (dotted), $r = 100$ (short dashed), $r = 150$ (long dashed), $r = 200$ (dot-dashed), and $r = 400$ (solid), respectively.

**Figure 3.** Nearest neighbor correlation for ISLs on generalized hypercubes as a function of the number of iterated smoothings. **a)** $\kappa = 2$, **b)** $\kappa = 4$.