Spectral Landscape Theory

Peter F. Stadler
Institut für Theoretische Chemie, Universität Wien
Währingerstrasse 17, A-1090 Wien, Austria
The Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM87501

The notion of an *adaptive landscape* has proved to be a valuable concept in theoretical investigations of evolutionary change, combinatorial optimization, and the physics of disordered systems. Landscape theory has emerged as an attempt to devise suitable mathematical structures for describing the "static" properties of landscapes as well as their influence on the dynamics of adaptation. Here we focus on the connections of landscape theory and algebraic combinatorics that form the basis of spectral approach to understanding landscape structure.

1 INTRODUCTION

Evolutionary change is caused by the spontaneously generated genetic variation and its subsequent fixation by drift and/or selection. Consequently, the main focus of evolutionary theory has been to understand the genetic structure and dynamics of populations, see e.g. [101]. In recent years, however, alternative approaches have gained increasing prominence in evolutionary theory. This development has been stimulated to some extent by the application of evolutionary models to designing *evolutionary algorithms* such as Genetic Al-

gorithms, Evolution Strategies, and Genetic Programming, as well as by the theory of Complex Adaptive Systems [69, 79, 38].

The generic structure of an evolutionary model is

$$x' = S(x, \mathbf{w}) \circ T(x, \mathbf{t}) \ , \tag{1}$$

where x is e.g. the vector of haplotype frequencies, and $S(x, \mathbf{w})$ is a term describing the selection forces acting on x. The parameters $\mathbf{w}$ form the so-called fitness function, since they can be regarded as a mapping from the set of types into the real numbers. The second term, $T(x, \mathbf{t})$, describes the transmission processes by determining the probability of transforming one type into another one by mutation and/or recombination [3]. Hence, evolution models can be seen as dynamical systems of genotype frequencies which live on an algebraic structure [96] that is determined by genetic processes such as mutation and recombination.

Metaphorically, the dynamics of evolutionary adaptation therefore can be seen as a walk on a landscape, where uphill moves are preferred. The realization that the topological features of fitness landscapes crucially influence the time-course of natural and simulated evolution led to what is now called *landscape theory*. It has several roots: In evolutionary theory it can be traced back to Wright's ideas about adaptive landscapes, see [114, pp. 304-317], and became important in theories of molecular evolution and the origin of life [31, 33, 42, 48, 75, 80, 110, 126, 127] and in evolutionary computer science [76, 77]. Similar developments exist in physics [44], where free energy landscapes of disordered systems such as spin glasses are considered [98], and in search theory [109]. The main challenge to landscape theory is to determine which features of the fitness landscape determine the evolvability of the systems on the landscape.

From the mathematical point of view, a landscape consists of three ingredients: (i) a set V of "configurations" which we shall assume to be finite but very large, (ii) a cost or fitness function $f : V \to \mathbb{R}$ that evaluates the configurations, and (iii) some sort of additional geometrical, topological, or algebraic structure $\mathcal{X}$ on V that allows us to define notions of closeness, similarity, or dissimilarity among the configurations. The structure $\mathcal{X}$, which turns the set V into the *configuration space* $(V, \mathcal{X})$, is determined by the particular application, e.g. a heuristic search procedure for a combinatorial optimization problem, or by the mechanisms of mutation and recombination in biological evolution.

A very promising approach in landscape theory is the decomposition of the fitness function $f : V \to \mathbb{R}$ is terms of a basis (of the vector space $\mathbb{R}^V$) that is induced in some natural way by $\mathcal{X}$. In other words, we search for a suitable spectral theory of the combinatorial space $(V, \mathcal{X})$, which we then use to "Fourier transform" f with respect to a suitable set eigenfunctions of $(V, \mathcal{X})$. The resulting "Fourier coefficients", so one hopes, will reveal the important features of the landscape much more readily than f itself.

2 CONFIGURATION SPACES

2.1 GRAPHS, HYPERGRAPHS, P-STRUCTURES, AND FINITE TOPOLOGIES

We have argued in the introduction that the "additional structure" $\mathcal{X}$ makes the boring "bag of numbers" $f : V \to \mathbb{R}$ a landscape – and an interesting mathematical object. The structure of the set V is oftentimes related to, or derived from, the internal structure of the objects $x \in V$. In this section we shall explore a few possibilities of imposing structure onto the set V of configurations.

2.1.1 Graphs. The simplest case is based on the notion of a *move set*. For each $x \in V$ we define as set $\mathcal{N}(x)$ of "neighbors" of x. The elements of $\mathcal{N}(x)$ are those configurations that can be reached in a single step starting form x. It will be convenient to assume $x \notin \mathcal{N}(x)$ for all x and to define $\overline{\mathcal{N}}(x) = \mathcal{N}(x) \cup \{x\}$. This definition allows us to regard the set

$$\mathcal{E} = \{ (x, y) \,|\, x \in V, y \in \mathcal{N}(x) \} \tag{2}$$

as the edge set of a directed graph with vertex set V. Equivalently, $\mathcal{E}$ is a *neighborhood relation* on V, that is, a relation satisfying $(x, x) \notin \mathcal{E}$ for all $x \in V$.

In many cases one is interested in symmetric neighborhood relations, i.e., in move sets in which each step is "reversible" and $(x, y) \in \mathcal{E}$ implies $(y, x) \in \mathcal{E}$. We may then regard V as an undirected graph with edges $\{x, y\} \in E$ iff $(x, y) \in \mathcal{E}$. The undirected graph case is by far the best studied one.

The tours of a traveling salesman problem, for example, can be encoded as the list of cities in the order in which they are visited. In other words, a particular tour is a permutation π of the cities $\{1, \ldots, n\}$. It seems natural to make use of the fact that these permutations form the symmetric group S_n: choose a move set $\Omega \subset \mathsf{S}_n$ and define that y is a neighbor if x if y is obtained from x by multiplication with an element of $t \in \Omega$, $y = xt$. Of course, we require that Ω does not contain the group identity $\imath$. Thus $(x, y) \in \mathcal{E}$ if and only if $x^{-1}y \in \Omega$. The resulting graph is a so-called *Cayley digraph* of the S_n. In most cases one assumes that $t \in \Omega$ implies $t^{-1} \in \Omega$, in which case the neighborhood relation is symmetric and the Cayley graph has undirected edges.

In molecular biology, we may for instance consider sequences as configurations and mutation as the move set. We have to distinguish two types of mutation: point mutations change a letter in a sequence without affecting its length. Insertions and deletions, on the other hand, change the length of a sequence. Editing operations such insertions and deletions are used for example in sequence alignment algorithms and can be generalized to trees [103, 146, 66]. Tree editing also provides a suitable analogue for mutation in Genetic Programming, see [106, 105]. Tree editing procedures are furthermore

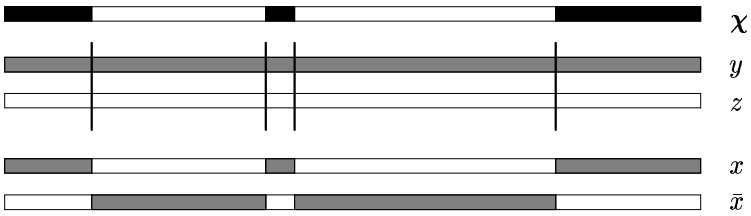

FIGURE 1 String Recombination. The children x and $\bar{x}$ are obtained by a multi-point crossover between the parents y and z. The bar on the top marks the positions that are transmitted together from a parent to an offspring in black and white, respectively.

used in phylogenetic reconstruction [35]. Editing operations related to sorting are used to deal with with genomic rearrangements [81].

2.1.2 Recombination, P-Structures, and Hypergraphs. Recombination, or cross-over, is another way of imposing a sense of closeness on V. For strings, the meaning of cross-over is easily defined.

A *cross-over operator* is a map $\chi : V \times V \to V \times V$ with the following property: Suppose $\chi(y, z) = (u, v)$. Then for each k either $y_k = u_k \wedge z_k = v_k$ or $z_k = u_k \wedge y_k = v_k$. By abuse of notation we write $x \in \chi(y, z)$ if $x = u$ or $x = v$, i.e., if x is an *offspring* of (y, z). As an immediate consequence we see that $\chi(x, x) = (x, x)$. We follow here the spirit of [77] when we regard a crossover operator as producing pairs of sequences rather than a single sequence from a pair of "ancestors".

We write $\boldsymbol{\chi} = \{k | x_k = u_k \wedge y_k = v_k\}$ and $\bar{\boldsymbol{\chi}} = \{k | y_k = u_k \wedge x_k = v_k\}$ for the two subsets of loci (sequence positions) that are *separated* by the cross-over operator χ. There is of course a one-to-one correspondence between a cross-over operator $\chi : V \times V \to V \times V$ and the associated set $\boldsymbol{\chi} \subseteq \{1, \ldots, n\}$: the set lists exactly those loci that are inherited from the first parent x by the first child u, see Figure 1. Analogously, $\bar{\boldsymbol{\chi}}$ is the list of loci that the first child u inherits from the second parent y. Of course, $\bar{\boldsymbol{\chi}} = \{1, \ldots, n\} \setminus \boldsymbol{\chi}$.

Note that ancestors and offsprings have the same length. More general types of recombination, often called *unequal crossover*, do not adhere to this restriction: here recombination evens may occur between independent positions in the two sequences whence the chain lengths of the recombinants vary [130]. Crossover operators may also be defined for trees and permutations, with applications in genetic programming [88] and GAs for the traveling salesman problem [159], respectively.

A *recombination operator* (in the sense of much of the GA literature) is a family $\mathcal{F}$ of cross-over operators that act on $V \times V$ with probabilities $\pi(\chi)$. In the following we restrict ourselves to equal-length crossover on strings. The

two most important recombination operators are uniform recombination [∞], consisting of all 2^n possible crossover operators, and 1-point recombination [1], which contains all cross-over operators χ for which the characteristic set is of the form $\chi = \{1, \ldots, k\}$.

Let V be a finite set with power set $\mathcal{P}(V)$. A *P-structure* [143] is a pair $(V, \mathcal{R})$ where $\mathcal{R} : V \times V \to \mathcal{P}(V)$. We say that the P-structure is *symmetric* if $\mathcal{R}(x, y) = \mathcal{R}(y, x)$ for all $x, y \in V$. In a weighted P-structure we attach a positive weight $\mathbf{H}_{x,(y,z)}$ to each triple (x, y, z) for which $x \in \mathcal{R}(y, z)$ and we set $\mathbf{H}_{x,(y,z)} = 0$ if $x \notin \mathcal{R}(y, z)$. We call $\mathbf{H}$ the incidence matrix of the P-structure.

In particular, there is a weighted P-structure associated in a natural way with each cross-over operator χ:

$$\begin{aligned} \mathcal{R}^\chi(y,z) &= \{x \in V \,|\, x \in \chi(y,z)\} \\ \mathbf{H}^\chi_{x,(y,z)} &= \begin{cases} 2 & \text{if} \quad x = y = z \\ 1 & \text{if} \quad x \in \chi(y,z) \qquad \text{and } y \neq z \\ 0 & \text{otherwise}. \end{cases} \end{aligned} \tag{3}$$

We observe that $\mathbf{H}^\chi_{x,(y,z)} > 0$ if and only if x is an offspring of (y, z). The doubled weight in the "diagonal", $\mathbf{H}^\chi_{x,(x,x)} = 2$, is mostly a technical convenience: It implies immediately $\sum_x \mathbf{H}_{x,(y,z)} = 2$, since any crossover operator produces exactly two offsprings from a pair of parents. If $y = z$, we simply count the offspring $y = z$ twice. The weighted P-structure associated with a recombination operator $\mathcal{F}$ is then

$$\begin{aligned} \mathbf{H} &= \sum_{\chi \in \mathcal{F}} \pi(\chi) \mathbf{H}^\chi \\ \mathcal{R}(y,z) &= \bigcup_{\chi \in \mathcal{F}} \mathcal{R}^\chi(y,z) = \left\{ x \in V \,\middle|\, \exists \chi \in \mathcal{F} \,:\, x \in \chi(y,z) \right\} \end{aligned} \tag{4}$$

The interpretation of this definition is straight forward: $\mathbf{H}_{x,(y,z)}$ is the chance that x is an offspring of the parents y and z under $\mathcal{F}$-recombination [153].

The *recombination hypergraph* imag$\mathcal{R}$ has vertex set V and hyper-edges $\mathcal{R}(y, z)$, $y, z \in V$. A spectral theory of hypergraphs is described in [123]. Gitchoff and Wagner [49] introduced a set axioms to describe the action of recombination in terms of what we call here P-structures. In [143, Lemma C2] we showed that any recombination operator forms a recombination structure if and only if the identity map on $V \times V$ is a member of the family $\mathcal{F}$ of crossover operators.

2.1.3 Finite Topological Spaces. The shape space of RNA secondary structure has been treated as a finite metric space, with a distance measure that is based on "structure editing" [39, 74]. It has become apparent, however, that distance measures of this type are not useful for explaining the features of evolutionary trajectories [40, 41]. In these contributions, a notion of "continuity"

is introduced and the evolutionary transitions are classified as continuous or discontinuous based on how easily one shape can be accessed from a previous one. Continuity is a topological property. Taking this idea serious one may regard shape space as (finite) topological space assuming that the "natural" topology is obtained by declaring the sets $\overline{\mathcal{N}}(x)$ of structures that are accessible from x as open sets. This approach will be pursued elsewhere [20]. We just mention here that finite topological spaces have a unique non-redundant basis consisting of the sets

$$\mathcal{B}(x) = \bigcap_{y:x\in\overline{\mathcal{N}}(y)} \overline{\mathcal{N}(y)} \tag{5}$$

which may be translated into the directed graph Υ with vertex set V and edges $\mathcal{B}(x) \setminus \{x\}$, $x \in V$. Topological properties such as separation properties can be then expressed as graph-theoretical properties of Υ, and we are back to the graph case.

2.2 MATRIX REPRESENTATIONS

2.2.1 Markov Chains. Not surprisingly we shall encounter a close relationship between spectral graph theory [12, 22, 21, 23] and landscapes on graphs in the course of this survey. A graph is faithfully represented by its adjacency matrix $\mathbf{A}$ which has the entries

$$\mathbf{A}_{xy} = \begin{cases} 1 & \text{if} \quad (x,y) \in \mathcal{E} \\ 0 & \text{if} \quad (x,y) \notin \mathcal{E} \end{cases} \tag{6}$$

Of course, $\mathbf{A}$ is symmetric if and only if the graph is undirected.

The most straight forward way of search on a possibly weighted (di)graph is a random walk, that is a Markov process with state space V. The most natural transition matrix is

$$\text{Prob}(y \to x) = \mathbf{S}_{xy} = \mathbf{A}_{xy} \Big/ \sum_{z\in V} \mathbf{A}_{xz}. \tag{7}$$

Such a random walk is usually called *simple* since each edge leaving y is chosen with the same probability. The denominator in equ.(7) is the out-degree of vertex x. The matrix $\mathbf{S}$ is the transition matrix of the random walk. Note that this is the transpose of the convention in most of the literature on Markov chains, see e.g. [11, 95]. The most important feature of random walks is the existence of a stationary distribution $\wp$ such that $\wp = \mathbf{S}\wp$ to which all initial distributions converge under fairly general conditions.

A Markov process is called reversible if its stationary distribution $\wp$ satisfies the *balance equation* $\mathbf{S}_{xy}\wp(y) = \wp(x)\mathbf{S}_{xy}$. In particular, a simple random walk on an undirected graph is reversible. Let $\mathbf{P}$ be the diagonal matrix with diagonal $\wp$. If $\mathbf{S}$ is the transition matrix of a reversible chain then

$\mathbf{T} = \mathbf{P}^{-1/2}\mathbf{S}\mathbf{P}^{1/2}$ is a bistochastic symmetric matrix. The "regularized" transition matrix $\mathbf{T}$ still essentially describes the graph Γ since $\mathbf{T}_{xy} > 0$ if and only if (x, y) is an edge of Γ. Since $\mathbf{T}$ is a symmetric non-negative matrix it serves as the starting point for the spectral theory of Markov processes [11].

Let us briefly consider the case of Hamming graphs in its most general setting. The configuration space consists of "genomes" with n loci (or positions) $k = 1, \ldots, n$. The are α_k alleles (or letters) at each position, which we denote by $x_k \in \mathcal{A}_k = \{0, 1, \ldots, \alpha_k - 1\}$. With $x_k \in \mathcal{A}_k$ we associate the root of unity

$$\hat{x}_k = \exp(2\pi\imath\, x_k/\alpha_k) \tag{8}$$

Furthermore, for $I \in \prod_k \mathcal{A}_k$ set $\tilde{I} = \{k | I_k \neq 0\}$. For each "index" I we define the *generalized Walsh function*

$$\varepsilon_I : \prod_k \mathcal{A}_k \to \mathbb{C} : \varepsilon_I(x) = \prod_k \hat{x}_k^{I_k} = \prod_{k\in\tilde{I}} \hat{x}_k^{I_k} \tag{9}$$

We remark that $\{\varepsilon_I | I \in V\}$ is the standard Fourier basis of the Abelian group $\prod Z_{\alpha_k}$, see section 2.3.6 below. These functions are eigenvectors of the adjacency matrix of the Hamming graphs $\prod \mathcal{Q}_{\alpha_k}$, i.e., the graphs obtained by considering point-mutations; see e.g. [135]. Note that the formal association of the index sets I with the vertices in V is a mere book-keeping device. If the number of alleles is the same for all loci, $\alpha_k = \alpha$, then the eigenvalue of $\mathbf{S}$ associated with ε_I is $\lambda_I = 1 - \frac{\alpha}{\alpha-1}|\tilde{I}|/n$.

Useful Markov processes on V can be defined, however, without any reference to a graph structure. The string recombination structures introduced in section 2.1.2 may serve as an example: A *cross-over walk* [70, 71] on V is the Markov process based on the following rule: The "father" y is mated with a randomly chosen "mother" z. The offsprings are the "son" x and the "daughter" $\bar{x}$. The "son" x becomes the "father" of the next mating. We regard the sequence of "fathers" as a random walk on V. It is straightforward to derive the transition matrix of this Markov process:

$$\mathbf{S}_{xy} = \frac{1}{2}\sum_{z\in V} \mathbf{H}_{x,(y,z)}\wp(z) \tag{10}$$

The factor $1/2$ stems from the fact that the offspring x is the "son" and not the "daughter" of the parents y and z with probability $1/2$. By $\wp(z)$ we denote the probability that z is the "mother" of the mating, i.e., $\wp(z)$ is the frequency of genotype z in the *population* $\wp$ under random mating. The uniform population case $\wp(z) = 1/|V|$, which is discussed in [143, 153], is generalized to the *Wright manifold* $\mathcal{W} = \{\wp | \wp(z) = \prod_k p_k(z_k)\}$, where $p_k(a)$ denotes the frequency of allele a at locus k, in [142]. It is not hard to verify that that linkage equilibrium is maintained under recombination, i.e., that any $\wp \in \mathcal{W}$ is a stationary distribution of $\mathbf{S}$ as defined in equ.(10). This fact

was first proven for two alleles and arbitrary number of loci by Robbins [121] and for multiple alleles by Bennett [9].

The "population-weighted" Walsh functions

$$\psi_I^{\wp}(x) = \prod_{k \in \tilde{I}} \frac{1}{p_k(x_k)} \hat{x}_k^{I_k} \tag{11}$$

are defined for all $\wp \in \mathcal{W}$. For a uniform population they coincide with the generalized Walsh functions introduced in equ.(9). In [142] we show that the population weighted generalized Walsh function $\psi_I^{\wp}$ is a left eigenvector of $\mathbf{S}^{\chi,\wp}$ with eigenvalue

$$\lambda_I^{\chi} = \begin{cases} 1 & \text{if} \quad \tilde{I} = \emptyset \\ 1/2 & \text{if} \quad \emptyset \neq \tilde{I} \subseteq \boldsymbol{\chi} \text{ or } \emptyset \neq \tilde{I} \subseteq \bar{\boldsymbol{\chi}} \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

This observation not only sets that stage for a spectral analysis of recombination landscapes, it also shows that recombination and mutation on strings are compatible operations that can therefore be compared directly in a meaningful way. The close relation between Hamming graphs and recombination spaces was noted with different methods by various groups, e.g. [19, 49, 93], see also appendix A.

2.2.2 Schrödinger Operators and Graph Laplacians. Let Γ be a simple graph (without loops and multiple edges), and let a be a weight function on the edges of Γ, conveniently defined as $a : V \times V \to \mathbb{R}_0^+$ such that $a(x,y) = a(y,x) > 0$ if $\{x,y\} \in E$ and $a(x,y) = 0$ otherwise. We say that Γ is unweighted if $a(x,y) \in \{0,1\}$, i.e., iff $a(x,y) = \mathbf{A}_{xy}$. Furthermore, let $v : V \to \mathbb{R}$ be an arbitrary *potential*. The linear operator $\mathbf{H}$ defined by the action

$$\mathbf{H}f(x) = \sum_{y \sim x} a(x,y)\,[f(x) - f(y)] + v(x)f(x) \tag{13}$$

is a discrete *Schrödinger operator* associated with Γ [24]. This definition includes the transition matrices of random walks on graph discussed in the previous section. The quantity

$$\deg(x) = -\sum_{y:\{x,y\} \in E} \mathbf{H}_{xy}$$

is the (generalized) *degree* of a vertex $x \in V$. The degree matrix $\mathbf{D}$ is the diagonal matrix of the vertex degrees.

We call $-\boldsymbol{\Delta} = \mathbf{D} - \mathbf{A}$ the Laplacian of the edge-weighted graph Γ [99, 100]. A slightly different definition is explored in [17]. Hence any Schrödinger operator is of the form $\mathbf{H} = -\boldsymbol{\Delta} + \mathbf{diag}(v(x))$. The Laplacian is therefore a Schrödinger operator without potential.

The analogy between discrete and continuous Schrödinger operators is a close one because the discrete Laplacian $-\mathbf{\Delta}$ resembles the Laplacian differential operator Δ is many ways. To see this, one introduces an arbitrary *orientation* on Γ by choose one of the two vertices u or v of the edge$h = \{v, w\}$ as the "positive end" and the other one as the "negative end". The matrix

$$\mathbf{\nabla}^{+}_{xh} = \begin{cases} +\sqrt{a(x,y)} & \text{if } x \text{ is the positive end of } h = \{x,y\} \\ -\sqrt{a(x,y)} & \text{if } x \text{ is the negative end of } h = \{x,y\} \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

is called the (weighed) *incidence matrix of* Γ. The choice of the symbol $\mathbf{\nabla}$ is intentional. In fact, let $f : V \to \mathbb{R}$ be an arbitrary function. Then $(\mathbf{\nabla} f)(h) = \sqrt{a(v,w)}[f(v) - f(w)]$ where h is the edge $\{v, w\}$, and v is the positive end of the edge h. This is as close to a first derivative as one can get on a graph. Note that $1/\sqrt{a(v,w)}$ takes the role of the distance between the vertices v and w.

The discrete Laplacian $-\mathbf{\Delta}$ is symmetric, non-negative definite, and singular. The eigenvector $(1, \ldots, 1)$ belongs to the eigenvalue $\tilde{\Lambda}_0 = 0$. Λ_0 has multiplicity 1 if and only if Γ is connected. A few simple computations verify that $\mathbf{\Delta} = -\mathbf{\nabla}^{+}\mathbf{\nabla}$ and hence corresponds to "second derivatives" on Γ. Let $\langle ., . \rangle$ denote the standard scalar product on $\mathbb{R}^{|V|}$, and let $f, g\, V \to \mathbb{R}$ be arbitrary landscapes. Then *Green's formula* holds in the following form:

$$\langle \mathbf{\nabla} f, \mathbf{\nabla} g \rangle = -\langle f, \mathbf{\Delta} g \rangle = -\langle g, \mathbf{\Delta} f \rangle \tag{15}$$

Graph Laplacians appear in very diverse fields of pure and applied mathematics. Their earliest use goes back to Kirchhoff's theory of electrical networks [84], see, e.g., [12, Chap.5].

2.2.3 Courant's Nodal Domain Theorem. A well-known feature of Schrödinger operators on Riemannian manifolds is that the nodal domains, that is, the connected components of $M \setminus \psi^{-1}(0)$, of their eigenfunctions are severely constrained. In order to formulate *Courant's theorem* for graphs, we define for any function $f : V \to \mathbb{R}$ on Γ: $\mathrm{supp}_{+}(f) = \{x \in V | f(x) > 0\}$, $\mathrm{supp}_{-}(f) = \{x \in V | f(x) > 0\}$, $\mathrm{zero}(f) = \{x \in V | f(x) > 0\}$, $\mathrm{supp}^{0}_{+}(f) = \mathrm{supp}_{+}(f) \cup \mathrm{zero}(f)$, and $\mathrm{supp}^{0}_{-}(f) = \mathrm{supp}_{-}(f) \cup \mathrm{zero}(f)$. A (strong) *nodal domain* of f is a maximal connected component of either $\mathrm{supp}_{+}(f)$ or $\mathrm{supp}_{-}(f)$. A *weak nodal domain* is a maximal connected component of $\mathrm{supp}_{+}(f) \cup \mathrm{zero}(f)$ or $\mathrm{supp}_{-}(f) \cup \mathrm{zero}(f)$, respectively.

Let $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{|V|}$ be the eigenvalues of a Schrödinger operator on Γ with corresponding eigenvectors φ_i. Define $M(i) = \max\{k | \lambda_k = \lambda_i\}$ and $m(i) = \min\{k \colon \lambda_k = \lambda_i\}$. Hence, $m(i) \leq i \leq M(i)$, $M(i) = m(i) + \mathrm{mult}(\lambda_i) - 1$, and $m(i) = M(i) = i$ if and only if λ_i is a simple eigenvalue of $\mathbf{H}$.

The main result on discrete Schrödinger operators is the following version of *Courant's Nodal Domain Theorem*, which motivates why the eigenfunctions of a Laplacian form particularly interesting basis sets for our purposes:

Let ψ_i be an eigenvector of $\mathbf{H}$ with eigenvalue λ_i. Then:
(i) There are at most $M(i)$ (strong) nodal domains of ψ.
(ii) There are at most $m(i)$ weak nodal domains of ψ.
(iii) If ψ_i has $m(i)+k$, $k>0$, (strong) nodal domains, then no two of them meet at a non-vertex point of the geometric representation of the graph Γ, and every vertex meets at least $k+1$ (strong) nodal domains.

The proofs of these results were obtained independently by different authors [45, 24, 149], beginning with Fiedler [36] who showed that the number of components of $\mathrm{supp}^0_+(\psi_i)$ is at most $M(i)$. Some closely related results on the component structure of $\mathrm{supp}_+(\psi_i) \cup \mathrm{supp}_-(\psi_i)$ can be found in [112].

2.3 SYMMETRIES, PARTITIONS, AND MATRIX ALGEBRAS

In many cases of practical interest there is a substantial amount of symmetry in the ways by which the set of configurations is constructed. Below we shall briefly explore a few approaches which haven been used to exploit these regularities in search of a workable spectral theory.

2.3.1 Relations and Automorphisms. A relation μ on V is simply a subset $\mu \subseteq V \times V$. The adjacency relation of a graph Γ may serve as an example. An *automorphism* of μ is a permutation $\mathsf{g} \in \mathsf{S}_{|V|}$ such that $(x,y) \in \mu$ if and only if $(\mathsf{g}(x), \mathsf{g}(y)) \in \mu$. The automorphisms of μ for the (permutation) group $\mathsf{Aut}[\mu]$, the *automorphism group* of μ. The automorphism group of a set $\mathcal{R}$ of relations on V is

$$\mathsf{Aut}[\mathcal{R}] = \bigcap_{\mu \in \mathcal{R}} \mathsf{Aut}[\mu] \tag{16}$$

To each relation μ on V there is an associated characteristic $|V| \times |V|$ matrix $\mathbf{R}^{(\mu)}$ with entries $\mathbf{R}^{(\mu)}_{xy} = 1$ if $(x,y) \in \mu$ and $\mathbf{R}^{(\mu)}_{xy} = 0$ if $(x,y) \notin \mu$.

Let G be an arbitrary permutation group acting on V. By $\mathsf{2orb}(\mathsf{G}, V)$ we denote the set of orbits of G acting on $V \times V$. Of course the $\mathsf{2orb}(\mathsf{G}, V)$ corresponds to a partition of $V \times V$, and each element of $\mathsf{2orb}$ may be regarded as a relation on V. These relation encapsulate the information about the symmetries that are most relevant for us.

A *matrix representation* of a finite group G is a map $\boldsymbol{\rho}$ from G into the group of $d \times d$ invertible matrices with complex coefficients such that $\boldsymbol{\rho}(\mathsf{gh}) = \boldsymbol{\rho}(\mathsf{g})\boldsymbol{\rho}(\mathsf{h})$ for all g, h in G. The *permutation representation* $\mathbf{G}$ of (G, V) consists of the $|V| \times |V|$ permutation matrices $\mathbf{G}(\mathsf{g})$ whose non-zero entries are $\mathbf{G}_{xy}(\mathsf{g}) = 1$ if and only if $x = \mathsf{g}(y)$.

A permutation group G on V is intimately connected with its *centralizer algebra*

$$\mathfrak{V} = \mathfrak{V}_{\mathbb{C}}(\mathsf{G}, V) = \left\{ \mathbf{M} \in \mathbb{C}^{|V|\times|V|} \,\middle|\, \forall \mathsf{g} \in \mathsf{G} : \mathbf{M}\mathbf{G}(\mathsf{g}) = \mathbf{G}(\mathsf{g})\mathbf{M} \right\} \tag{17}$$

The set $\mathfrak{V}$ is closed w.r.t. addition and multiplication of matrices and w.r.t. multiplication with scalars from the underlying field $\mathbb{C}$. Its dimension (as vec-

tor space) equals the rank of its permutation group, $\dim(\mathfrak{V}) = \mathrm{rank}(\mathsf{G}, V)$. The characteristic matrices $\mathbf{R}^{(\mu)}$ of the orbits $\mu \in \mathsf{2orb}(\mathsf{G}, V)$ form the *standard basis* of the vector space $\mathfrak{V}$. From $\mathbf{R}^{(\mu)} \circ \mathbf{R}^{(\nu)} = \delta_{\mu,\nu}\mathbf{R}^{\mu}$ we see that $\mathfrak{V}$ is also closed under component-wise (Schur or Hadamard) multiplication. Finally, $\mathfrak{V}$ is closed under transposition since the transpose $\mu^+ = \{(x,y)|(y,x) \in \mu\}$ of an orbit is again an orbit.

2.3.2 Coherent Algebras. A set of complex matrices that is closed under (i) scalar multiplication with complex numbers, (ii) component-wise addition, (iii) ordinary matrix multiplication, (iv) component-wise multiplication, and (v) transposition is called a *coherent algebra* or *cellular algebra*. Equivalently, a matrix algebra $\mathfrak{W} \subseteq \mathbb{C}^{|V|\times|V|}$ is coherent if and only if it satisfies the following axioms:
(i) As a linear space over $\mathbb{C}$, $\mathfrak{W}$ has a basis of $\{\mathbf{R}^{(1)}, \ldots, \mathbf{R}^{(r)}\}$ of 0-1 matrices.
(ii) $\sum_{j=1}^{n} \mathbf{R}^{(j)} = \mathbf{J}$, the all-1 matrix.
(iii) For every $i \in \{1, \ldots, r\}$ there an i' such that $\mathbf{R}^{(i)\mathsf{T}} = \mathbf{R}^{(i')}$.
(iv) $\mathbf{I} \in \mathfrak{W}$.
Sometimes coherent algebras without unity are considered, i.e., axiom (iv) is disregarded. The centralizer algebras of permutation groups form the most prominent class of coherent algebras (with identity).

Axiom (ii) above implies that the relations associated with the basis matrices $\mathbf{R}^{(j)}$ form a partition of $V \times V$. Such partitions are known as *coherent configurations* [63, 64, 65]. Table 1 gives an overview of various properties of partitions of $V \times V$ that are of interest in the context of landscapes. For details see e.g. [135, 136].

For each collection $\mathcal{M} = \{\mathbf{M}_1, \ldots, \mathbf{M}_k\}$ of $|V| \times |V|$ matrices there is a smallest coherent algebra $\langle\langle \mathcal{M} \rangle\rangle$ which is the defined as the intersection of all coherent algebras that contain $\{\mathbf{M}_1, \ldots, \mathbf{M}_k\}$. Since the centralizer algebra is coherent we have

$$\langle\langle \mathcal{M} \rangle\rangle \subseteq \mathfrak{V}_{\mathbb{C}}(\mathsf{Aut}[\mathcal{M}], V) \tag{18}$$

Equality hold if and only if there is a permutation group that has $\langle\langle \mathcal{M} \rangle\rangle$ as its centralizer algebra [87]. The coherent algebra $\langle\langle \mathcal{M} \rangle\rangle$ can therefore be regarded as a "combinatorial approximation" of the centralizer algebra [34, 85]. This is of particular importance in the graph case: given the adjacency matrix $\mathbf{A}$ of Γ, there is polynomial time algorithm that determines the coherent algebra $\mathfrak{W}(\Gamma) = \langle\langle \mathbf{A} \rangle\rangle$, see [157, 7, 6].

Let $\mathcal{R} = \{\mathbf{R}^{(1)}, \ldots, \mathbf{R}^{(r)}\}$ be the standard basis of a coherent algebra $\mathfrak{W}$. We have $\mathbf{R}^{(\mu)}\mathbf{R}^{(\nu)} = p^{\kappa}_{\mu,\nu}\mathbf{R}^{(\kappa)}$ where *intersection numbers*

$$p^{\kappa}_{\mu,\nu} = \left|\left\{ z \in V \;\middle|\; (x,z) \in \mu \wedge (z,y) \in \nu \right\}\right| \in \mathbb{N}_0 \tag{19}$$

are the same for all pairs $(x,y) \in \kappa$. The $r \times r$ matrices $\hat{\mathbf{R}}^{\kappa}$ with entries $\hat{\mathbf{R}}^{(\nu)}_{\mu,\kappa} = p^{\kappa}_{\mu,\nu}$ generate a matrix algebra $\hat{\mathfrak{W}}$ that is isomorphic to $\mathfrak{W}$ [63]. This

TABLE 1 Regularity properties of partitions.

i	$\mu \cap \mathcal{I} \neq \emptyset \Longrightarrow \mu \subseteq \mathcal{I}$ for any $\mu \in \mathcal{R}$
ii	$\mathcal{I} \in \mathcal{R}$
iii	$\{x \mid \exists y \in V : (x,y) \in \mu\} = \{y \mid \exists x \in V : (x,y) \in \mu\} = V$ for all $\mu \in \mathcal{R}$
iv	$\mu \in \mathcal{R} \Longrightarrow \mu^+ \in \mathcal{R}$
v	$\mu = \mu^{\mathsf{T}}$ for all $\mu \in \mathcal{R}$
vi	$\left\lvert\{y \in V \mid (x,y) \in \mu\}\right\rvert$ and $\left\lvert\{y \in V \mid (y,x) \in \mu\}\right\rvert$ depend only on $\mu \in \mathcal{R}$ but not on $x \in V$
vii	The numbers $p^{\kappa}_{\mu,\nu} = \left\lvert\left\{ z \in V \,\middle\vert\, (x,z) \in \mu \wedge (z,y) \in \nu \right\}\right\rvert$ are the same for all pairs $(x,y) \in \kappa$
viii	The matrices $\mathbf{R}^{(\mu)}$ and $\mathbf{R}^{(\nu)}$ commute for all $\mu, \nu \in \mathcal{R}$

Property	i	ii	iii	iv	v	vi	vii	viii
homogeneous	∘	•						
transitive	•	∘	•					
precoherent configuration	∘	•		•				
symmetric					•			
class degree regular			∘			•		
homogeneous class deg. regular	•	∘	∘			•		
coherent	•			•			•	
homogeneous coherent configuration	∘	•	∘	•		∘	•	
class degree regular cc	•	∘	∘	•		•	•	
association scheme	•	∘	∘	•		∘	•	•
symmetric association scheme	•	∘	∘	∘	•	∘	•	∘

$\mathcal{I} = \{(x,x) \mid x \in V\}$ is called the diagonal of $V \times V$. The symbol • indicates properties that are used for definition, while ∘ marks additional properties that are implied by the definition.

observation makes coherent algebras appealing objects for our purposes because $\hat{\mathfrak{W}}$ is small enough in many cases to allows for explicit computations.

The action of a permutation group (G, V) is *transitive* if for all $x, y \in V$ there is $\mathsf{g} \in \mathsf{G}$ such that $y = \mathsf{g}(x)$. If $\mathsf{g} \in \mathsf{G}$ can be chosen such that $y = \mathsf{g}(x)$ and $x = \mathsf{g}(y)$ then (G, V) is *generously transitive*. Generously transitive permutations groups have symmetric (and therefore commutative) centralizer algebras see [63] and [160, Thm. 29.3]. The group case is summarized in Table 2.

2.3.3 Association Schemes. If the coherent algebra $\mathfrak{W}$ with standard basis $\mathcal{R} = \{\mathbf{R}^{(1)}, \ldots, \mathbf{R}^{(r)}\}$ is commutative, we obtain a symmetric association scheme $\mathfrak{W}^{\sigma}$ by taking as basis elements $\mathbf{R}^{(\mu)}$ if $\mu = \mu^{\mathsf{T}}$ and $\mathbf{R}^{(\mu)} + \mathbf{R}^{(\mu^{\mathsf{T}})}$ if $\mu \neq \mu^{\mathsf{T}}$, respectively. In particular, if a graph has a generously transitive or at least multiplicity-free group of automorphisms then $\mathfrak{W}(\Gamma) = \langle\langle \mathbf{A} \rangle\rangle$ is a (symmetric) association scheme.

TABLE 2 Permutation Groups and Their Centralizer Algebras

G		$\mathfrak{V}_{\mathbb{C}}(\mathsf{G}, V)$
transitive	$\Longleftrightarrow$	homogeneous
multiplicity free	$\Longleftrightarrow$	commutative
generously transitive	$\Longleftrightarrow$	symmetric

The situation becomes particularly simple in this case. Since $\mathfrak{W}$ is a commutative algebra of symmetric matrices, a so-called *Bose-Mesner algebra* [14, 25], there is a common basis $\mathbf{\Phi} = \{\varphi_i : V \to \mathbb{C},\ i = 1, \ldots, |V|\}$ of eigenvectors of all matrices $\mathbf{M} \in \mathfrak{W}$.

The following observation is also of interest in this context [136, Lemma 9]: The coherent algebra $\mathfrak{W}[\Gamma]$ of a graph is always a refinement of the *distance partition* of Γ which has the classes $\delta_d = \{(x, y) \in V | d(x, y) = d\}$, i.e., $\mathbf{A}^{(d)} \in \langle\langle \mathbf{A} \rangle\rangle$, where $\mathbf{A}^{(d)}$ is the characteristic matrix of δ_d. The class of *distance regular graphs*, which contains important examples such as the Hamming graphs $\mathcal{Q}_\alpha^n$ and the Johnson graphs, is characterized by the fact that the distance partition forms a (symmetric) association scheme. These graphs have received considerable attention, see e.g. [16].

2.3.4 Adjacency Algebra and Hoffman Algebras The adjacency algebra of a graph Γ is the matrix algebra generated by the adjacency matrix, $\mathfrak{A}[\Gamma] = \langle \mathbf{A} \rangle$. Clearly, $\mathfrak{A}[\Gamma] \subseteq \mathfrak{W}[\Gamma]$. Higman [64] showed that $\mathfrak{A}[\Gamma] = \mathfrak{W}[\Gamma]$ if and only if $\mathfrak{W}[\Gamma]$ is commutative, i.e., an association scheme. It is interesting to note in this context that a homogeneous coherent algebra with rank $r \leq 5$ is always commutative [64]. So-called "orbit-polynomial graphs" characterized by $\mathfrak{A}[\Gamma] = \mathfrak{V}[\mathsf{Aut}[\Gamma], V]$ are considered in [8].

A *Hoffman algebra* is matrix algebra $\mathfrak{H} \subseteq \mathbb{C}^{|V| \times |V|}$ such that (i) there is a basis consisting of non-negative integer matrices and (ii) $\mathbf{J} \in \mathfrak{H}$ [86]. This notion is of interest as a generalization of coherent algebras and since $\mathfrak{A}[\Gamma]$ is a Hoffman algebra, i.e., $\mathbf{J} \in \mathfrak{A}[\Gamma]$ if and only if Γ is a connected regular graph [67].

2.3.5 Equitable Partitions. Consider the set $\mathcal{R}$ of relations associated with the coherent algebra $\mathfrak{W}[\Gamma]$ of a graph Γ. Fix $x_0 \in V$ and define

$$\mu[x_0] = \{y \in V | (x, x_0) \in \mu\} \tag{20}$$

Clearly, $\Pi(x_0) = \{\mu[x_0] | \mu \in \mathcal{R}\}$ is a partition of V. If $\mathfrak{W}$ is homogeneous, then $\mu[x_0]$ is non-empty for all $\mu \in \mathcal{R}$, see Table 1, i.e., $\mathcal{R}$ and $\Pi[x_0]$ have the

same number of classes for all $x_0 \in V$. In [64] it is shown that

$$\hat{\mathbf{R}}^{(\nu)}_{\mu\kappa} = p^{\kappa}_{\mu,\nu} = \sum_{x\in\nu[x_0]} \hat{\mathbf{R}}^{(\nu)}_{xz} \qquad \text{for each } z \in \kappa[x_0] \tag{21}$$

Noting that the adjacency matrix $\hat{\mathbf{A}}$ of Γ is a sum of $\mathbf{R}^{(\mu)}$-matrices, it is shown in [136] that

$$\sum_{x\in\mu[x_0]} \mathbf{A}_{xy} = \tilde{\mathbf{A}}_{\mu[x_0],\nu[x_0]} = \sum_{\kappa\subseteq\mathcal{E}} p^{\kappa}_{\mu,\nu} = \hat{\mathbf{A}}_{\mu,\nu} \tag{22}$$

holds for any $y \in \nu[x_0]$. Partitions of V that satisfy the first equality in equation (22) are called *equitable*. Equitable partitions have been introduced by Schwenk [128]; more recently they have been used by Powers and coworkers as "colorations", see, e.g., [113, 111]. In [22, Chap.4] they appear as "divisors" of graphs.

The most important property of an equitable partition Π is that all eigenvalues of the *collapsed adjacency matrix* $\tilde{\mathbf{A}}$ are also eigenvalues of $\mathbf{A}$. If Π contains a class that consists of single vertex, then the minimal polynomials of $\tilde{\mathbf{A}}$ and $\mathbf{A}$ are the same [13, Thm.8.6], i.e., the relevant spectral information is already contained in $\tilde{\mathbf{A}}$. More information about equitable partitions can be found in [50, 51, 52].

2.3.6 Fourier Transform on Finite Groups and Cayley Graphs. Let $\mathbf{G}$ be a finite group and let $f : \mathsf{G} \to \mathbb{C}$. Let $\boldsymbol{\rho}$ be matrix representation of G. Then

$$\hat{f}(\boldsymbol{\rho}) = \frac{1}{\sqrt{|\mathsf{G}|}} \sum_{\mathrm{g}\in\mathsf{G}} f(\mathrm{g})\boldsymbol{\rho}(\mathrm{g}) \tag{23}$$

is called the *Fourier transform*[1] *of* f *at* $\boldsymbol{\rho}$. The Fourier transform on a complete set $\mathcal{R}$ of irreducible representations is inverted by

$$f(\mathrm{g}) = \frac{1}{\sqrt{|\mathsf{G}|}} \sum_{\boldsymbol{\rho}\in\mathcal{R}} \dim\boldsymbol{\rho}\, \mathrm{Tr}\left[\hat{f}(\boldsymbol{\rho})\boldsymbol{\rho}(\mathrm{g}^{-1})\right] \tag{24}$$

Fast Fourier Transform algorithms are known for a variety of finite groups. For a recent overview see e.g. [97, 122].

It is not surprising that the spectral properties of Cayley graphs are intimately related to the Fourier transform on the underlying group. The crucial observation is the following. Let δ_Ω be the characteristic function of the set of generators Ω. Then $\sqrt{|\mathsf{G}|}\hat{\delta}_\Omega(\boldsymbol{\rho}_{\mathrm{reg}})$, the Fourier transform of δ_Ω at the regular representation of G equals the adjacency matrix of $\Gamma(\mathsf{G}, \Omega)$ up to a reordering of the group elements. Its spectrum is therefore the union of the spectra of $\sqrt{|\mathsf{G}|}\hat{\delta}_\Omega(\boldsymbol{\rho}_i)$ where $\boldsymbol{\rho}_i$ are the irreducible representations of G. If Ω is a union of conjugacy classes of G the situation simplifies further [29].

[1] In most of the literature the normalization factor $|\mathsf{G}|^{-1/2}$ is omitted

The irreducible representations are all 1-dimensional if G commutative. Since G can be written as a direct product of cyclic groups, $\mathsf{G} = \prod_{k=1}^{m} \mathsf{C}_{n_k}$, the characters are

$$\chi_{\mathsf{g}}(\mathsf{x}) = \exp\left(2\pi i \sum_{k=1}^{m} \frac{x_k g_k}{n_k}\right) \tag{25}$$

where we use the additive representation of C_{n_k} as $\{0, 1, \ldots, n_k - 1\}$ with addition modulo n_k. It is not hard to verify that the characters χ_{g} are eigenvectors of the adjacency matrix of each Cayley graph of G. The corresponding eigenvalue are $\sum_{\mathsf{x}\in\Omega} \chi_{\mathsf{g}}(\mathsf{x})$, see e.g. [94]. The Fourier transform on C_2^n is also known as the *Walsh-Hadamard* transform. Note the that the Boolean hypercube can be regarded as a Cayley graph on this group. An FFT algorithm for this case is due to Yates [163]. Further material about the Cayley graphs on commutative groups can be found in [1].

3 LANDSCAPES

3.1 FOURIER DECOMPOSITION AND ELEMENTARY LANDSCAPES

Having derived a set of basis functions $\{\varphi_k | V \to \mathbb{C}\}$ from the structure of a configuration space $(V, \mathcal{X})$ by means of one of the approaches outlines in the previous section, it is natural to expand the fitness function f in terms of this basis:

$$f(x) = \sum_k a_k \varphi_k(x) \tag{26}$$

We shall use the following convention: (i) The index 0 is reserved for the "ground state". If the basis is derived from a Laplacian, for instance, then φ_0 is constant, the associated eigenvalue is zero, and

$$a_0 = \sum_x \varphi_0(x) f(x) = |V|^{-1} \sum_{x\in V} f(x) \tag{27}$$

Similarly, the index 0 will refer to the stationary distribution in the case of a Markov chain on V. (ii) The distinct eigenvalues of $-\mathbf{\Delta}$ will be denoted by Λ_p, in the Markov chain case we write λ_p. It will be convenient to define the index sets $J_p = \{k | -\mathbf{\Delta}\varphi_k = \Lambda_p \varphi_k\}$ that collect all eigenfunctions belonging to the same (Laplacian) eigenvalue. (iii) We write $\tilde{f}(x) = f(x) - a_0$. If φ_0 is constant, this is the "non-flat" part of fitness function.

Lov Grover and others [18, 57, 136] observed that $\tilde{f}$ is in many cases an eigenfunction of the graph Laplacian $-\mathbf{\Delta}$, see Table 3 for a list of examples. We say that f is *elementary* w.r.t. $-\mathbf{\Delta}$ if $\tilde{f}$ is an eigenfunction of $-\mathbf{\Delta}$ with an eigenvalue $\lambda_p < 1$. In [142] this notion is extended to calling f elementary

TABLE 3 Elementary Landscapes

Problem	Graph	D	λ	State
p-spin glass	$\mathcal{Q}_2^n$	n	$2p$	p
NAES$^{(1)}$	$\mathcal{Q}_2^n$	n	4	2
Weight Partitioning	$\mathcal{Q}_2^n$	n	4	2
Graph α-Coloring	$\mathcal{Q}_2^\alpha$	$(\alpha-1)/n$	2α	2
XY-spin glass	$\mathcal{Q}_\alpha^n$	$(\alpha-1)/n$		2
for $\alpha > 2$:	$\mathcal{C}_\alpha^n$	2		2
TSP symmetric	$\Gamma(\mathcal{S}_n, \mathcal{T})$	$n(n-1)/2$	$2(n-1)$	2
	$\Gamma(\mathcal{S}_n, \mathcal{J})$	$n(n-1)/2$	n	2
	$\Gamma(\mathcal{A}_n, \mathcal{C}_3)$	$n(n-1)(n-2)/6$	$(n-1)(n-2)$	?
antisymmetric	$\Gamma(\mathcal{S}_n, \mathcal{T})$	$n(n-1)/2$	$2n$	3
	$\Gamma(\mathcal{S}_n, \mathcal{J})$	$n(n-1)/2$	$n(n+1)/2$	$\mathcal{O}(n)$
Graph Matching	$\Gamma(\mathcal{S}_n, \mathcal{T})$	$n(n-1)/2$	$2(n-1)$	2
Graph Bipartitioning	$J(n, n/2)$	$n^2/4$	$2(n-1)$	2

w.r.t. a random walk transition operator iff $\mathbf{S}\tilde{f} = \lambda_p \tilde{f}$ with an eigenvalue $\lambda_p < 1$.

If f is elementary, then $\tilde{f}$ satisfies the conditions of Courant's nodal domain theorem, see 2.2.3. Elementary landscapes thus can be expected to have few nodal domains if they belong to a small Laplacian eigenvalue (or to an eigenvalue of Markov transition matrix close to 1), while landscapes that are far away from the ground state will in general have many nodal domains. Such landscapes will appear "rugged". Grover [57] showed that

$$f(\hat{x}_{\min}) \leq a_0 \leq f(\hat{x}_{\max}) \tag{28}$$

where $\hat{x}_{\min}$ and $x_{\max}$ are arbitrary local minima and maxima, respectively. This *maximum principle* shows that elementary landscapes are well-behaved: There are no local optima with worse than average fitness. We shall return to local optima as a measure of ruggedness in section 4.4.

In section 2.2.1 we have seen that p-spin (or Walsh-) functions are the eigenfunctions not only of mutation operators but also of recombination operators. Indeed, there is an intriguing relationship between elementary landscapes for string recombination and schemata *sensu* Holland [2, 10, 68, 69], see also Appendix A. Each recombination-elementary landscape corresponds to a partitioning of the set of strings. Each equivalence class in this partitioning is a schema in the sense of Holland and all the schema which make up this partitioning have the same positions fixed. An elementary landscape in this context is a landscape which assumes that only the fixed positions in the schema actually influence fitness. This was first noted by Weinberger in his seminal paper on Fourier and Taylor series of fitness landscapes [155]. In

[143, 153] it is shown rigorously that this is a legitimate way of decomposing the configuration space of string recombination.

3.2 CORRELATION MEASURES

3.2.1 Random Walk Autocorrelation Functions The ruggedness of a landscape is most easily quantified by measuring the correlation of fitness values in "neighboring" positions. Weinberger [154, 155] suggested the following procedure. Given a Markov process on V, we sample the fitness values $f(x^{(t)})$, interpret them as a time series, and compute the autocorrelation function of this time series. Let $\mathbf{T}$ be the transition matrix of a such reversible Markov process with stationary distribution φ_0. We define the scalar product

$$\langle f, g\rangle_{\varphi_0} = \sum_{x\in V} f(x)\varphi_0(x)g^*(x) \tag{29}$$

where a^* denotes the complex conjugate of a. The (expected) autocorrelation function along a $\mathbf{T}$-random walk on V is then

$$r(t) = \left(\sum_{x\in V} |\tilde{f}^2(x)|\varphi_0(x)\right)^{-1} \sum_{y\in V} \tilde{f}(x)(\mathbf{T}^t)_{xy}\tilde{f}^*(y)\varphi_0(y) = \frac{\langle \tilde{f}, \mathbf{T}^t\tilde{f}\rangle_{\varphi_0}}{\langle \tilde{f}, \tilde{f}\rangle_{\varphi_0}} \tag{30}$$

Expanding f w.r.t. eigenvectors of $\mathbf{T}$ it can be shown [136] that

$$r(t) = \sum_{p\neq 0} B_p\lambda_p^t \qquad \text{with} \qquad B_p = \frac{\sum_{i\in J_p}|a_i|^2}{\sum_{i\neq 0}|a_i|^2}\ . \tag{31}$$

Thus a landscape f is elementary w.r.t. a transition operator $\mathbf{T}$ if and only if the "random walk" autocorrelation function is exponential, $r(t) = \lambda_p^t$. In this case the the *order* p indicates to which eigenvalue (not counting multiplicities) $\tilde{f}$ belongs. On a Boolean hypercube $\mathcal{Q}_\alpha^2$ we have eigenfunctions of the form $\sum_I a_I \prod_{k\in I} x_k$ where $p = |I|$ is constant. These are exactly Derrida's [26] p-spin models. The order of the elementary landscape thus equals the "interaction order" of the underlying spin glass model.

3.2.2 Amplitude Spectra Equ.(31) decomposes non-elementary landscapes in a natural way into a superposition of elementary ones. The amplitudes B_p measure the relative variance contributions of the different eigenspaces (or "modes"). Instead of the random walk correlation function $r(t)$ we can therefore use the *amplitude spectrum* B_p, $p \geq 1$, as a measure for the ruggedness of a landscape. In many cases it is much easier to interpret than the correlation function, see e.g. [136, 46, 72]. This technique was applied successfully to realistic landscapes such as those arising from RNA folding, see Figure 2 for an example. The RNA secondary structure folding model is described in detail in Peter Schuster's contribution to this book.

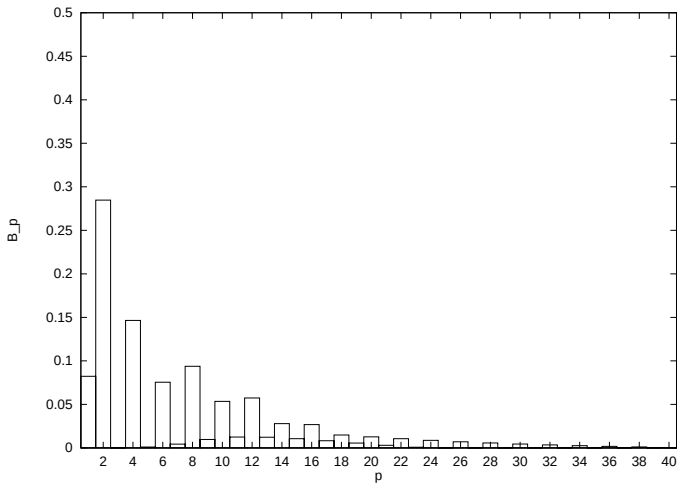

FIGURE 2 The estimated amplitude spectrum for a GC landscape with $n = 100$ under mutation [72]. The configuration space is the Hamming graph $\mathcal{Q}_2^{100}$ of sequences taken from the 2-letter alphabet {G,C}.
The most striking feature of the amplitude spectrum of RNA landscapes is a strong difference between even and odd modes. This can easily be explained in terms of the physics underlying RNA folding: The major contribution of the folding energy comes from stacking of base pairs. Hence the major changes in free energy caused by a point mutation will arise from these contributions. Since stacking energies are influenced by even number of nucleotides depending on the location of the affected base pair within a stack. A recent comparison of amplitude spectra for different landscapes based on folding short RNA chains indicates that the amplitude spectra of the free energy landscapes are typical [137].

3.2.3 Distance Correlation Functions Most "early" work on RNA landscapes, e.g. [42, 145] uses a different type of correlation measure based on the Hamming distance. In [136, 139] a more general version starting with a collection of relations on V is introduced. Let μ be a relation on V. Then we set

$$\varrho(\mu) = \frac{|V|^2}{|\mu|} \frac{\sum_{(x,y)\in\mu}(f(x)-\bar{f})(f(y)-\bar{f})}{\sum_{x,y\in V}(f(x)-\bar{f})(f(y)-\bar{f})} \tag{32}$$

where $\bar{f} = |V|^{-1}\sum_x f(x)$. Thus $\varrho(\mu)$ is the variation of points of vertices within a relation μ compared to the variance of f over all configurations $x \in V$. On Hamming graphs, for instance, it is natural to consider the distance classes, i.e., $(x,y) \in \mu_d$ if and only if $d_H(x,y) = d$, a predefined value. Such distance-dependent correlation functions have been considered also for some combinatorial optimization problems [4, 5, 131, 138]. Given a partition of $V \times V$, we may of course regard ϱ as a function of the classes of this parti-

tion. Furthermore, if this partition is sufficiently "nice", then the correlation function ϱ itself also has useful algebraic properties. The main result of [134], for instance, is the following theorem: Let f be landscape on a regular graph Γ that has a homogeneous coherent algebra $\mathfrak{W}[\Gamma]$. Then $r(s)$ is exponential if and only if ϱ is a left eigenvector of the collapsed adjacency matrix $\hat{\mathbf{A}}$.

4 RANDOM LANDSCAPES

4.1 BASIC DEFINITIONS

In many cases, for instance in applications to spin glasses, the definition of the landscape contains a number random parameters. We therefore define *random landscapes* as elements of an appropriate probability space, following the presentation in [118].

Let V be a finite set and let W be a predicate of landscapes $f : V \to \mathbb{R}$. A *random W-landscape* over V is the probability space

$$\mathbf{\Omega} = (\{f : V \to \mathbb{R} \mid f \text{ has property } W\}, \mathcal{A}, \mu) \ , \tag{33}$$

where $\mathcal{A}$ is a σ-field and $\mu : \mathcal{A} \to [0,1]$ a measure. Let $\xi : \mathbf{\Omega} \to \mathbb{R}$ be an Ω-random variable; we denote expectation value and variance of ξ by $\mathbb{E}[\xi]$ and $\mathbb{V}[\xi]$, respectively. In particular, we shall write $\mathbb{E}[f(x)]$ for the expected value of $f(x)$, i.e., of $\mathbb{E}[f]$ evaluated at $x \in V$.

The *covariance matrix* $\mathbf{C}$ of the random landscape $\mathbf{\Omega}$, is given componentwise by

$$\mathbf{C}_{xy} = \mathbb{E}[f(x)f(y)] - \mathbb{E}[f(x)]\mathbb{E}[f(y)] \tag{34}$$

Clearly, $\mathbf{C}$ is a symmetric non-negative definite $|V| \times |V|$ matrix. Taking the set of *all* maps, $\{f : V \to \mathbb{R}\}$, as basis space of the probability space $\mathbf{\Omega}$, a basis is formed by a set of orthonormal eigenvectors $\{\psi_k\}$ of the covariance matrix $\mathbf{C}$. An expansion of the form

$$f(x) \doteq \sum_k b_k \psi_k(x) \tag{35}$$

is know as *Karhunen-Loève series* or *principal component decomposition*. The symbol $\doteq$ denotes equality almost surely. The importance of equ.(35) comes from the following classical result [73]: Then the coefficients $\{b_k\}$ in equ.(35) are uncorrelated random variables satisfying

$$\mathrm{Cov}[b_k, b_j] = \theta_k \delta_{kj}, \qquad 1 \le k, l \le |V| \, , \tag{36}$$

where $\theta_k = \mathbb{V}[b_k]$ is the eigenvalue of $\mathbf{C}$ belonging to the eigenvector ψ_k.

A random landscape is elementary if $\mathbb{E}[r(t)] = \lambda_p^s$ or, equivalently, if $\mathbb{E}[B_p] > 0$ only for a single mode $p > 0$. While Ising spin glasses, TSPs with

random coefficients and other random parameter variants of combinatorial optimization problems are elementary, this is not true for Kauffman's Nk models [139]. As a consequence, there are many landscapes that cannot be constructed as a superposition of Nk-models, for instance the Sherrington-Kirkpatrick spin glass [129], see [135, 60, 61] for a detailed discussion of the amplitude spectrum $\mathbb{E}[B_p]$ of Nk models.

A random landscape is *pseudo-isotropic* [139] if there are constants a_0, v, and w such that for all $x \in V$ holds (i) $\mathbb{E}[f(x)] = a_0$, (ii) $\mathbb{V}[f(x)] = v^2$, and (iii) $|V|^{-1} \sum_{y \in V} \mathbf{C}_{xy} = w$. Pseudo-isotropy is a fairly weak regularity property that is satisfied by many random landscape models of practical importance, see Table 4.

4.2 ADDITIVE RANDOM LANDSCAPES

Many important random landscapes can be written as a sum of components with random coefficients. More precisely, let M be finite index set, let c_j, $j \in M$ be independent, real valued random variables over appropriate probability spaces $\Omega_j = (\mathbb{R}, \mathcal{A}_j, \mu_j)$, and let $\Theta = \{\vartheta_j : V \to \mathbb{R} \,|\, j \in M\}$ be a family of real valued functions on V. An *additive random landscape* (arl) is the probability space $(\mathbf{\Omega}_V, \otimes_j \mathcal{A}_j, \otimes_j, \mu_j)$ with

$$\mathbf{\Omega}_V = \left\{ f : V \longrightarrow \mathbb{R} \mid f(x) = \sum_{j=1}^{M} c_j \vartheta_j(x) \right\} . \tag{37}$$

In other words, the random landscape is constructed as a linear combination of non-random landscapes ϑ_j with independent random coefficients c_j.

In particular, any Gaussian random landscape is additive: Using the Karhunen-Loève decomposition, equ.(35), any random landscape can be written as linear combination with uncorrelated random coefficients; uncorrelated Gaussian random variables are independent.

The most important additive random landscapes exhibit further regularities: An arl is *uniform* if and only if (i) the random variables c_i, $i \in M$, are i.i.d. and (ii) there exist constants $a, b \in \mathbb{R}$ such that $\sum_{x \in V} \vartheta_i(x) = |V| a$ and $\sum_{x \in V} \vartheta_i^2(x) = |V| b$. A uniform random landscape is *strictly uniform* if there exist constants $d, e \in \mathbb{R}$ such that $\sum_j \vartheta_i(x) = d$ and $\sum_j \vartheta_i^2(x) = e$. In [118] we show that a uniform random landscape is pseudo-isotropic if and only if at least one of the following two conditions is satisfied: (i) $\mathcal{F}$ is strictly uniform, or (ii) $a = 0$, $\mathbb{E}[c_i] = 0$, and there is a constant $e \in \mathbb{R}$ such $\sum_i \vartheta_i^2(x) = e$ for all $x \in V$.

4.3 ISOTROPY AS A MAXIMUM ENTROPY CONDITION

Uniformity and pseudo-isotropy are still rather weak properties. In [135, 139] the notion of an *isotropic* random landscape was introduced as a "statistically symmetric model", that is, as a random landscape with a covariance matrix

TABLE 4 Examples of Additive Random Landscapes.
The component landscapes ϑ_I and the index set M, equ.(37), are listed together with information whether the models are uniform (U), strictly uniform (S), or pseudo-isotropic (P). As in Table 1, properties that are implied by stronger ones are shown as ∘.

Model	Component Landscapes and Index set	U	S	P
Ising spin glass	$\vartheta_I(x) = \prod_{k\in I} x_k \qquad I \subseteq \{1,\dots,n\}$	•		•
SK model	as above with $\lvert I\rvert = 2$	•		•
Nk Landscapes	see [118]	∘	•	∘
Graph Bipartitioning	$\vartheta_{ij}([A,B]) = \begin{cases} 1 & \text{if } \{i,j\} \not\subseteq A,B \\ 0 & \text{otherwise} \end{cases} \quad i<j$	∘	•	∘
Asymmetric TSP	$\vartheta_{kl}(\tau) = \sum_i \delta_{k,\tau(i)}\delta_{l,\tau(i-1)} \qquad k \neq l$	∘	•	∘

that shares the symmetries of the underlying configuration space. More precisely, a random landscape is *isotropic w.r.t. a partition* $\mathcal{R}$ of $V \times V$ if there are constants a_0 and s and a function $c : \mathcal{R} \to \mathbb{R}$ such that (i) $\mathbb{E}[f](x) = a_0$ and $\mathbb{V}[f](x) = s^2$ for all $x \in V$, and (ii) $\mathbf{C}_{xy} = c(\mu)$ for all $(x,y) \in \mu$, i.e., the covariance matrix $\mathbf{C}$ is constant on the classes $\mu \in \mathcal{R}$.

The notion of isotropy for random landscapes is the analogue of *stationarity* for stochastic processes. Following the conventions of Karlin and Taylor [78] our notion of isotropy would be called "covariance isotropic", "weakly isotropic", or "wide sense isotropic". For a Gaussian random landscape the notions of (weak) isotropy and strict isotropy coincide of course.

Not surprisingly, a useful theory does not arise by considering arbitrary partitions $\mathcal{R}$, see Table 1. Isotropy is a stronger concept than pseudo-isotropy only if $\mathcal{R}$ is sufficiently regular. Transitivity, for instance, ensures that the classes of $\mathcal{R}$ are large enough to be interesting. In [139, thm.4] the following result is proved: Let $\mathcal{F}$ be isotropic w.r.t. a homogeneous class degree regular partition $\mathcal{R}$ of $V \times V$. Then $\mathcal{F}$ is pseudo-isotropic. Furthermore, suppose $\mathbb{E}[f(x)] = a_0$ for all $x \in V$. Then the random landscape is isotropic w.r.t. a homogeneous coherent configuration if and only if $\mathbf{C} \in \langle\langle \mathcal{R} \rangle\rangle$ [139].

If $\mathbf{A}$ is the adjacency matrix of an undirected graph (or more generally, the a symmetric transition matrix of a Markov process on V then we say that a random landscape is **-isotropic w.r.t.* $\mathbf{A}$ if $\mathbb{E}[f(x)] = a_0$ and $\mathbf{C} \in \langle \mathbf{A} \rangle$. For association schemes (such as those arising from distance regular graphs including the hypercube) isotropy and *-isotropy equivalent. In [139] we show that a random landscape is *-isotropic if and only the Fourier coefficients (w.r.t. an orthonormal basis of eigenvectors of $\mathbf{A}$) satisfy: (i) $\mathbb{E}[a_k] = 0$ for $k \neq 0$, (ii) $\mathrm{Cov}[a_k, a_j] = \delta_{kj}\mathbb{V}[a_k]$, and (iii) $\mathbb{V}[a_k] = \mathbb{V}[a_j]$ if $k,j \in J_p$. These conditions mean that the Fourier coefficients are uncorrelated and that they have the same mean and variance whenever they belong to the same mode

(eigenspace of $\mathbf{A}$). Hence Fourier and Karhunen-Loève series coincide for *-isotropic landscapes.

For a random landscape with measure μ we define the *entropy*

$$S = -\int \mu(f) \ln \mu(f) df \tag{38}$$

In appendix B we review some well-known properties of the entropy functional. In particular, S can be decomposed into a "homogeneous" part S_{σ^2} that only depends on $\sigma^2 = \mathrm{Tr}\mathbf{C}$, the total variance of landscape, a terms $S_{\mathbf{C}}$ that depends only on the variations among the eigenvalues of $\mathbf{C}$, and a third term that measures the effect of deviations from the normal distribution given a fixed covariance matrix $\mathbf{C}$. Given $\mathbf{C}$, a random landscape maximizes entropy if and only if its Gaussian.

In the following we assume a Gaussian random landscape. Now suppose the values $\beta_p = \sum_{k \in I_p} \mathbb{V}[a_k]$ are prescribed, too. It follows from the discussion in Appendix B that the entropy is maximized if and only if the covariance matrix restricted to J_p is a multiple of the identity, i.e., iff $\mathbb{V}[a_k]$ if constant on J_p. Given its amplitude spectrum, a random landscape therefore maximized entropy if and only if it is Gaussian and *-isotropic. This is of practical interest since the class *-isotropic models (on their natural configuration spaces) includes among others Derrida's p-spin Hamiltonians, the graph-bipartitioning problem, and the TSP.

Most variants of Kauffman's Nk-model, the XY-Hamiltonians, short-range Ising models, or the Graph-Matching Problem are not isotropic. This has important implications for the structure of these landscapes, as we shall see below.

4.4 LOCAL OPTIMA

Palmer [108] used the existence of a large number of local optima to define ruggedness. We say that $x \in V$ is a *local minimum* of the landscape f if $f(x) \leq f(y)$ for all neighbors y of x. The use of $\leq$ instead of $<$ is conventional [82, 124]; it does not make a significant difference for spin glass models. Local maxima are defined analogously. The number $\mathcal{N}$ of local optima of a landscape, however, is much harder to determine than its autocorrelation function $r(s)$ or its correlation length

$$\ell = \sum_{k=0}^{\infty} r(s) = \sum_{p \neq 0} \frac{B_p}{1 - \lambda_p} \tag{39}$$

As it appears that $\mathcal{N}$ and ℓ are two sides of the same coin we search for a connection between the two quantities. In a random landscape setting it is customary to determine $\mathbb{E}[\ln \mathcal{N}]$. The only known case in which $\mathbb{E}[\ln \mathcal{N}] \neq \ln \mathbb{E}[\mathcal{N}]$ is the linear spin chain [27].

For the case of short range spin glasses, in which only a small number z of coupling constants J_{ij} are non-zero for any given spin i, a slightly larger number of local optima has been found [15, 147] than for the long-range Sherrington-Kirkpatrick model [129]. Since all Ising models have the same correlation length $\ell = n/4$ [156, 133] but somewhat different values of $\mathcal{N}$, we cannot hope for a general, exact formula relating $\mathbb{E}[\ln \mathcal{N}]$ and $\mathbb{E}[\ell]$.

From the maximum entropy interpretation of isotropy, however, we know that the expected density of meta-stable states in an isotropic Gaussian random landscape is determined completely by the expected correlation function $\mathbb{E}[r(s)]$ because such a model simply does not contain any further information. In the case of an elementary isotropic random landscape the correlation length ℓ already determines $r(s)$ and hence there must be a direct relationship between $\mathbb{E}[\ell]$ and the expected number of meta-stable states $\mathbb{E}[\ln \mathcal{N}]$. Its functional form will of course depend on the geometric properties of Γ.

Stadler and Schnabl [141] conjectured that $\mathbb{E}[\ln \mathcal{N}]$ can be estimated as follows: For a typical elementary landscape we expect that the correlation length ℓ gives a good description of its structure because the landscape does not have any other distinctive features. By construction ℓ determines the size of the mountains and valleys. As there are many directions available at each configuration we expect there are only very few meta-stable states besides the summit of each of these ℓ-sized mountains – almost all of the configurations will be saddle points with at least a few superior neighbors. We measure ℓ along a random walk but the radius $R(\ell)$ of a mountain is more conveniently described in terms of the distance between vertices on Γ. Here $R(\ell)$ is the average distance that is reached by the random walk in ℓ steps. With the notation $B(R)$ for the number of vertices contained in a ball of radius R in Γ we expect approximately $|V|/B(R(\ell))$ local optima.

As an example we consider a comparison of the correlation length conjecture with an exact computation based on the TAP equations [148] for Derrida's p-spin Hamiltonian [26]. The TAP approach yields [56, 120]:

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{n} \ln \mathbb{E}[\mathcal{N}] = \alpha(0) &= \ln 2 - \frac{\delta(p)^2}{2(p-1)} + \ln \Phi(\delta(p)) \qquad \text{with} \\ \Phi(x) &= (1 + \operatorname{erf}(\delta(p)/\sqrt{2}))/2 \\ \delta(p) &= \frac{p-1}{\sqrt{2\pi}} \frac{\exp(-\delta(p)^2/2)}{\Phi(\delta(p))} \end{aligned} \tag{40}$$

The last equation must be solved numerically for $\delta(p)$.

An explicit evaluation of the correlation length conjecture for the p-spin Hamiltonian [140] yields

$$\begin{aligned} \alpha(0) &= \ln 2 + (1-\zeta)\ln(1-\zeta) + \zeta \ln \zeta \qquad \text{where} \\ \zeta &= \frac{1}{2}\left(1 - \mathrm{e}^{-1/p}\right) . \end{aligned} \tag{41}$$

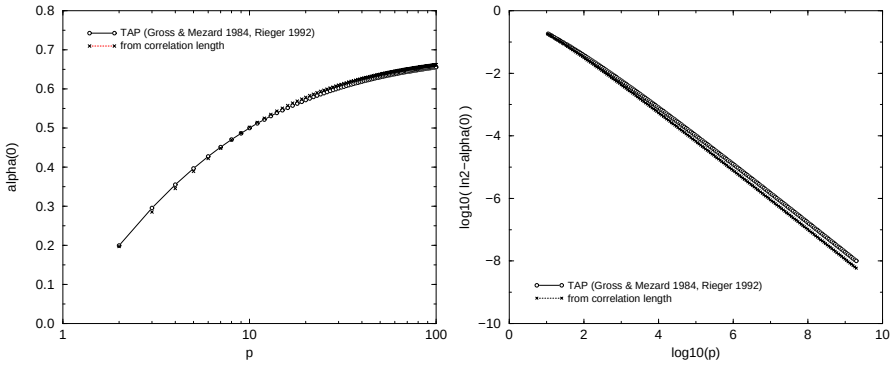

FIGURE 3 Comparison of eqns.(40) and (41) for small (l.h.s) and large (r.h.s.) values of p.

Equ.(41) compares very will with data numerical simulations for $p = 2, 3, \ldots, 6$ [140]. Figure 3 shows that equ.(41) is in excellent agreement with the TAP result in equ.(40). It is interesting to note that even for very large p there is a good qualitative agreement between the value of $\alpha^* = \ln 2 - \alpha(0)$ obtained by the two methods. We find $\alpha^*_{\mathrm{TAP}} \sim p^{-1} \ln p$ and $\alpha^*_{\mathrm{clc}} \sim (1/2p) \ln p$.

The correlation length conjecture works very well for isotropic random landscapes on various other configuration space besides Boolean hypercubes [141, 89]. On the other hand, the correlation length conjecture yields sometimes very poor estimates if the landscapes deviate significantly from isotropy [46].

4.5 NEUTRALITY

We say two configurations $x, y \in V$ are *neutral* if $f(x) = f(y)$. We colloquially refer to a landscape as "neutral" if a substantial fraction of adjacent pairs of configurations are neutral. This should not be confused with the *flat* landscape, in which f is constant. Extensive computer simulations, based on RNA secondary structures [125, 58, 59], have revealed that neutrality plays an important role in understanding the dynamics of RNA evolution [75, 119, 117].

Kimura proposed a theory of biological evolution that focuses exclusively on the aspects of neutrality [83] by assuming a flat fitness landscape. Very recently, landscapes with a large degree of neutrality have also been described in computational models such as cellular automata [71], for the mapping of sequences in combinatorial random structures [116], and in the context of sequential dynamical systems [90].

Neutral landscapes are discussed in more detail in Christian Reidys' contribution to this book. Here we restrict ourselves to indicating how a rigorous investigation of neutrality can be linked to the techniques described above. In the case of additive random landscapes, a promising starting point is provided

by the random variables $X_{\{y,x\}}(f)$ which take the value 1 if $f(x) = f(y)$ and 0 otherwise.

The number of *neutral neighbors* of a configuration $x \in V$ is then

$$\nu_x(f) = \sum_{y \in \mathcal{N}(x)} X_{\{y,x\}}(f) \tag{42}$$

The following parameters have turned out to be particularly important for understanding the neutrality in an additive random landscape:

$$\begin{aligned} c_x(y) &= \left|\{j \in M \mid \vartheta_j(x) \neq \vartheta_j(y)\}\right| \\ w_x(y',y'') &= \left|\{j \in M \mid \vartheta_j(x) \neq \vartheta_j(y') \ \wedge \vartheta_j(x) \neq \vartheta_j(y'')\}\right| \end{aligned} \tag{43}$$

where $x \in V$ is an arbitrary vertex and $y, y', y'' \in \mathcal{N}(x)$. Note that $c_x(y)$ and $w_x(y', y'')$ only depend on the properties of the component landscapes ϑ_k but not on the distribution of the coefficients c_k.

In [118] explicit expressions for the mean and variance of neutrality are derived for the simplest possible case, namely for a distribution of the coefficients c_j satisfying

$$\mu\{c_j = \xi\} = \begin{cases} \mu_0 > 0 & \text{if} \xi = 0 \\ 0 & \text{otherwise} \end{cases} \tag{44}$$

For any additive random landscape with coefficients c_i satisfying equ.(44) we obtain

$$\begin{aligned} \mathbb{E}[\nu_x] &= \sum_{y \in \mathcal{N}(x)} \mu_0^{c_x(y)} \\ \mathbb{V}[\nu_x] &= \sum_{y',y'' \in \mathcal{N}(x)} \mu_0^{c_x(y')+c_x(y'')} \left[\mu_0^{-w_x(y',y'')} - 1\right] \end{aligned} \tag{45}$$

More explicitly, the expected number of neutral neighbors of a p-spin landscape is therefore $\mathbb{E}[\nu] = n\ \mu_0^{\binom{n-1}{p-1}}$. Depending on μ_0, the expected fraction of vanishing interaction coefficients, the fraction of neutral mutations $\mathbb{E}[\nu]/n$ may take any value between 0 and 1. This fact is independent of the order p of the spin glass. Thus ruggedness (as measured by p) and neutrality (as measured by μ_0 are independent properties of (random) landscapes.

In many spin glass models the spins are arranged on a finite-dimensional lattice. Hence each spin has only a finite number of other spins to which it is couples in such a *short range spin glass*. All but $\mathcal{O}(n)$ coefficients therefore vanish and we have $\mu_0 \sim 1 - z/n^{p-1}$, where $z > 0$ is a parameter determined by the connectivity of the lattice. The fraction of neutral spin flips is constant in such systems, $\mathbb{E}[\nu]/n \sim \mathrm{e}^{-z}$, see [118] for more details. We remark further that the fraction of neutral mutations is the crucial input parameter for random graph models of neutral landscapes [47, 115, 119], see also Chr. Reidy's contribution to this book.

The study of neutrality in more general classes of random landscapes requires the determination of the distributions of the random variables

$$\delta(x,x') := \sum_{i:\,\vartheta_i(x)\neq\vartheta_i(x')} c_i\,[\vartheta_i(x)-\vartheta_i(x')] \tag{46}$$

Since the c_i are by definition independent in an additive landscape, we have to compute the convolutions

$$g_{x,x'}(\delta) = \underset{i:\vartheta_i(x)\neq\vartheta_i(x')}{\star} \rho_i\big(c_i/[\vartheta_i(x)-\vartheta_i(x')]\big) \tag{47}$$

where $\star_j$ denotes convolution of all functions indexed by j, $g_{x,x'}(\delta)$ is the density of the values of $\delta(x,x')$, and $\rho_i(.)$ is the density function of c_i, which in this case has to be evaluated with the argument $c_i/[\vartheta_i(x)\neq\vartheta_i(x')]$. Then we have

$$\mathrm{Prob}[X_{x'}(x)=1] = \lim_{\epsilon\to 0}\int_{-\epsilon}^{\epsilon} g_{x,x'}(\delta)d\delta\ . \tag{48}$$

If $\{i|\,\vartheta_i(x)\neq\vartheta_i(x')\}=\varnothing$, x and x' are neutral for any distribution of the c_i. This is observed for instance in the *graph matching problem* [132].

We conclude from equ.(48) that a continuous density $g_{x,x'}$, which necessarily arises if the individual densities ρ_i are continuous, does not lead to neutrality. Neutrality hence depends on a "discrete" contribution to the probability densities of the coefficients c_k, $k\in M$. Indeed, only these discrete components influence neutrality. In practice, evaluation of the convolution (47) therefore boils down to a combinatorial exercise, as for instance in the case of the integer-valued NK model proposed by Newman and Engelhardt [102].

Finally we remark that $g(\,.\,)$ would also be the appropriate starting point for a theory of *nearly neutral landscapes* [104], in which the condition such as $|f(x)-f(x')|<\varepsilon$ for some finite $\varepsilon>0$ could replace the condition $f(x)=f(x')$.

5 DISCUSSION

The exposition above has been focussed almost entirely on the "static" properties of a landscape. A mathematical language has been introduced that allows us to view a cost or fitness function as it is seen by a search operator. This formalism lends a precise meaning to notions such as ruggedness, neutrality, or isotropy. Intuitively, the dynamics of (evolutionary) adaptation and the performance of optimization heuristics should be determined by exactly these properties.

The spectral approach described here has been useful in distinguishing different types of landscapes. RNA folding landscapes, for instance, are very

different from spin glasses in both ruggedness and neutrality. A more detailed analysis of the distribution of the Fourier coefficients that belong to a particular mode might help to understand and quantify the structure of anisotropies.

Dynamics on landscapes, unfortunately, is much less understood at present. Apart from a few global results such as the "No Free Lunch Theorem" [162] and detailed studies on very simple landscapes such as [150], very few exact results are known. Various dynamical phenomena have been described for special classes of landscapes. There is an error-threshold limiting the mutation rate in biological evolution [32] which is well understood at least on landscapes with a few peaks. A tunnel effect was described between two separated peaks [30]. On the other hand, a diffusion-like process is observed on landscapes with a high degree of neutrality [75, 117] similar to the situation in a flat landscape [28]. A theory that could treat all these aspects within a common formalism, however, is still missing.

ACKNOWLEDGMENTS

Stimulating discussion with Jim Crutchfield, Jan Cupal, Walter Fontana, Silvio Franz, Ricardo García-Pelayo, Ivo Hofacker, Wim Hordijk, Stuart Kauffman, Richard Palmer, Christian Reidys, Peter Schuster, Gottfried Tinhofer, Henri Waelbroeck, Günter Wagner and many others made this work possible. Section 4.4 originated during a workshop at the ICTP in Trieste in 1997.

APPENDIX

A SCHEMATA AND DECEPTIVENESS

A.1 Introduction. Walsh functions and "schemata" have been used extensively in the analysis of GA behavior, see e.g. [53, 54, 55, 92, 107, 37, 151, 152]. A *schema* is simply a hyperplane in sequence space. It is defined by the set H of "fixed" bits and their values h_i, $i \in H$. In symbols

$$\mathcal{H} = H[h] = \{x \in V | \forall i \in H : x_i = h_i\}\,. \tag{49}$$

For a discussion of the Schema Theorem and the Building Block Hypothesis we refer to the literature [2, 10, 68, 43, 69, 144]. Instead, we briefly consider a few properties of landscapes that a naturally defined in terms of schemata. For simplicity we restrict ourselves to landscapes on the set of binary strings of length n. Notion such as "local optimum" in the following subsections consequently refer to the graph structure of the Boolean hypergraph. Schemata and Walsh functions are linked by means of

$$f(\mathcal{H}) = \sum_{I \subseteq H} a_I \varepsilon(h) \quad \text{and} \qquad \mathrm{var}(\mathcal{H}) = \sum_{I \subseteq H} \left\{ \sum_{K \not\subseteq H} a_K a_{K \triangle I} \right\} \varepsilon_I(h) \tag{50}$$

where $I \triangle J$ denotes the symmetric difference of the sets I and J. Note that these quantities are superpositions of Walsh functions with index set $I \subseteq H$ evaluated at the fixed bits of the schema. It is interesting that $f(\mathcal{H})$ depends only on Fourier coefficients a_I with $I \in H$, while $\mathrm{var}(\mathcal{H})$ depends only on coefficients with $I \not\subseteq H$.

Let Q be any property of a landscape f on V. We say that f is robustly Q, if there is an $\epsilon > 0$ such that any landscape g satisfying $|f(x) - g(x)| < \epsilon$ for all $x \in V$ also has property Q. The condition $|f(x) - g(x)| < \epsilon$ may be replaced by $|a_k - \tilde{a}_k| < \epsilon$ for the Fourier coefficients of f and g, respectively. The results in following sections have not been published before; nevertheless we omit their (rather simple) proofs in this survey.

A.2 Funnels A landscape is called a *funnel* if there is a string h^* such that $H \subseteq H'$ implies $f(H[h^*]) \leq f(H'[h^*])$ for all $H \subseteq \{1, \ldots, n\}$. If f is a funnel with peak h^* then (i) $f(h^*) \geq \bar{f}$ and (ii) h^* is a local maximum on $\mathcal{Q}_2^n$. It is not hard to construct simple examples of funnels with just $n = 3$ bits showing that the peak of a funnel need not be globally optimal, and that there may be multiple local optima. With some more work it is also possible to show that a landscape is a robust funnel if and only if $H \subset H'$ implies $f(H[h^*]) < f(H'[h^*])$ for all H. The peak h^* of a robust funnel is unique.

A.3 GA-Easy Functions Let us call a function *GA-easy* if there is a global optimum x^* such that $f(H[x^*]) \geq f(H[x])$ for all $x \in V$ and all $H \subseteq [n]$. In the GA literature a more common definition of easy is what we call *robustly GA-easy*, namely a function f with a global optimum x^* satisfying $f(H[x^*]) > f(H[x])$ for all $x \in V$ and all $H \subseteq [n]$ for which $H[x^*] \neq H[x]$. By setting $H = \{1, \ldots, n\}$ we see that the global optimum of a robustly GA-easy function is unique. Robustly GA-easy is called "fully easy" in [91]. We prefer to say that f is *fully GA-easy* if for each global optimum x^*, for each $H \subseteq [n]$, and for each $x \in V$ holds $f(H[x^*]) \geq f(H[x])$. Naturally, a fully GA-easy function is GA-easy, but the converse is not true. Note also that "robustly fully GA-easy" is the same as robustly GA-easy. A short computation shows that a (robustly) GA-easy function with global optimum x^* is a (robust) funnel with peak x^*.

Linear functions f are of course GA-easy. However, linear functions do not form a generic class of landscapes in the sense that linearity is not a robust property. Wilson [161] introduced a slightly larger class of landscapes: Given a string $x \in V$ let $\hat{x}$ be a string satisfying $\hat{x}_k = x_k$ if $f(x) > f(x^{(k)})$ and $\hat{x}_k = \bar{x}_k$ if $f(x) < f(x^{(k)})$. That is, $\hat{x}$ is obtained from x by keeping the best bit among all 1-error mutants in each position. A function f on V is *bit setting optimizable* (b.s.o.) if $\hat{x}$ is a global optimum for each $x \in V$. Of course linear functions are b.s.o., and the set of robustly bso functions is nonempty: all sufficiently small perturbations of linear functions are b.s.o. It can be shown that a (robustly) b.s.o. function is (robustly) GA-easy. It is shown in [161] that the converse is not true for all $n \geq 3$.

A.4 Deceptive Functions The literature on deceptive functions uses a variety of slightly different notions of deceptiveness (and sometimes does not even precisely define the notion at all). In a deceptive landscape an optimal schema of some size is "contradicted" by one of its sub-schemata. Intuitively, this is just the converse of GA-easy. Following Whitley [158] we use the following formal definition: A landscape f is *deceptive* if there are vertices $x, y \in V$ and index sets $H \subset K \subset \{1, \ldots, n\}$ with the following properties: (i) $K[x] \neq K[y]$, (ii) $f(H[x]) > f(H[z])$ for all $z \in V$ with $H[x] \neq H[z]$, and (iii) $f(K[y]) > f(K[z])$ for all $z \in V$ with $K[y] \neq K[z]$. As expected, it can be shown that a GA-easy function f is not deceptive. However, the converse it not true since a "symmetric function", i.e., a function fulfilling $f(x) = f(\bar{x})$ where $\bar{x}$ is the complement of x is never deceptive according to Whitley's definition. An example of a symmetric function that is not not GA-easy can be constructed e.g. on $\mathcal{Q}_2^4$. We will say that f is *GA-hard* if it is not GA easy.

We say that f is *weakly deceptive* if there are vertices $x, y \in V$ and index sets $H \subset K$ such that (i) $f(H[x]) \geq f(H[z])$ for all $z \in V$ and $f(H[x]) > f(H[y])$, and (ii) $f(K[y]) \geq f(K[z])$ for all $z \in V$ and $f(K[y]) > f(K[x])$. A deceptive landscape is of course weakly deceptive. Since symmetric functions can be weakly deceptive, the converse is not true in general.

Let Ω denote the set of global optima. For each $x^* \in \Omega$ and each index set H we define the set of vertices that belong to a H-schema that is superior to $H[x^*]$:

$$\Psi(x^*, H) = \left\{ y \middle| \, f(H[y]) > f(H[x^*]) \,\text{and}\, \forall\, z : f(H[y]) \geq f(H[z]) \right\} . \qquad (51)$$

Clearly $x^* \notin \Psi(x^*, H)$. The following propositions are easily verified:
(i) f is GA-hard iff for each $x^* \in \Omega$ there is an index set H such that $\Psi(x^*, H)$ is non-empty.
(ii) f is not fully GA-easy iff there is a $x^* \in \Omega$ and an index set H such that $\Psi(x^*, H)$ is non-empty.
(iii) f is weakly deceptive iff there is a $x^* \in \Omega$ and an index set H such that $\Psi(x^*, H) \setminus \Omega$ is non-empty.

Both weak deceptiveness and GA-hardness imply that f is not fully GA-easy. If f has a unique global optimum, however, then "weakly deceptive", "GA-hard", and not "fully GA-easy" are equivalent properties.

B: MAXIMUM ENTROPY CONDITIONS

It is well known the Gaussian distributions maximize entropy. The proof for the one-dimensional case can be found e.g. in [62, prop. 1.15]. For the convenience of the reader a short proof of the general case is included here as it is not readily accessible in the literature.

The starting point is the following inequality that holds for arbitrary probability spaces:

$$\int p(x) \ln p(x) dx - \int p(x) \ln q(x) dx \geq 0 \qquad (52)$$

Equality holds in equ.(52) holds if and only if $p = q$ almost everywhere. Let $\mathbf{C}$ be the covariance matrix of p. We assume that $\mathbf{C}$ is invertible. Without loosing generality we furthermore assume $\mathbb{E}[x] = 0$. Substituting the Gaussian distribution

$$q(x) = \frac{1}{(2\pi)^{|V|/2}\sqrt{\det \mathbf{C}}} \exp\left(-\frac{1}{2} x \mathbf{C}^{-1} x\right) \tag{53}$$

equ.(52) translates into a general inequality for the the entropy of p:

$$S \leq \frac{|V|}{2} \ln(2\pi) + \frac{1}{2} \ln \det \mathbf{C} + \frac{1}{2} \int_{\mathbb{R}^n} (x \mathbf{C}^{-1} x) p(x) dx$$

The integral is a simple constant independent of p as the following computation shows

$$\int_{\mathbb{R}^n} \sum_{k,l} (\mathbf{C}^{-1})_{kl} x_k x_l p(x) dx = \sum_{k,l} (\mathbf{C}^{-1})_{kl} \int_{\mathbb{R}^n} x_k x_l p(x) dx =$$
$$\sum_{k,l} (\mathbf{C}^{-1})_{kl} \mathbf{C}_{kl} = \sum_k (\mathbf{C}^{-1}\mathbf{C})_{kk} = \mathrm{Tr}\mathbf{I} = |V|.$$

Let $\{\Lambda_k,\, k = 1, \ldots, |V|\}$ be the eigenvalues of $\mathbf{C}$. Since $\mathbf{C}$ is invertible by assumption we have $\Lambda_k > 0$ for all k. Using $\sigma^2 = \mathrm{Tr}\mathbf{C} = \sum_k \Lambda_k$ we obtain

$$S \leq S_{\mathbf{C}} = \frac{1}{2}|V| \ln \frac{2\pi e}{|V|} + \frac{1}{2} \sum_k \ln \frac{\Lambda_k |V|}{\sigma^2} \tag{54}$$

It is easy to verify that $S_{\mathbf{C}}$ is indeed the entropy of a Gaussian distribution with covariance matrix $\mathbf{C}$.

The two terms in equ.(54) allow for a direct interpretation. The Gaussian entropy $S_{\mathbf{C}}$ attains its maximum subject to a given variance σ^2 if and only if $\Lambda_k = \sigma^2/|V|$, in which case the second term vanishes. We may this split the entropy of a random landscape into three contributions

$$S = S_{\sigma^2} + \Delta S_{\mathbf{C}} + \Delta S_{ng} \tag{55}$$

where $\Delta S_{ng} = S - S_{\mathbf{C}}$ is the entropy loss due to deviations from a Gaussian distribution, S_{σ^2} is the maximal entropy with given variance σ^2, and $\Delta S_{\mathbf{C}}$, the second term in equ.(54), measures the entropy loss due to variations in the spectrum of $\mathbf{C}$. In particular, whenever there are correlations between different vertices, then $\mathbf{C}$ is non-diagonal and hence $\Delta S_{\mathbf{C}} < 0$. More precisely, $\Delta S_{\mathbf{C}} = 0$ if and only if the corresponding Gaussian random landscape is i.i.d.

REFERENCES

[1] B. Alspach. Isomorphism and Cayley graphs on abelian groups. In G. Hahn and G. Sabidussi, editors, *Graph Symmetry: Algebraic Methods and Applications*, volume 497 of *NATO ASI Series C*, pages 1–22. Kluwer, Dordrecht, 1997.

[2] L. Altenberg. The schema theorem and the Price's theorem. In L. D. Whitley and M. D. Vose, editors, *Foundations of Genetic Algorithms 3*, pages 23–49, San Francisco CA, 1995. Morgan Kauffman.

[3] L. Altenberg and M. W. Feldman. Selection, generalized transmission, and the evolution of modifier genes. I. The reduction principle. *Genetics*, 117:559–572, 1987.

[4] E. Angel and V. Zissimopoulos. Autocorrelation coefficient for the graph bipartitioning problem. *Theor. Computer Sci.*, 191:229–243, 1998.

[5] E. Angel and V. Zissimopoulos. On the quality of local search for the quadratic assignment problem. *Discr. Appl. Math.*, 82:15–25, 1998.

[6] L. Babel, S. Baumann, M. Lüdecke, and G. Tinhofer. `STABCOL`: Graph isomorphism testing based on the Weisfeiler-Leman algorithm. Technical Report TUM-M9702, TU München, Garching, Germany, 1997.

[7] L. Babel, I. V. Chuvaeva, M. Klin, and D. V. Pasechnik. Algebraic combinatorics in mathematical chemistry. Methods and algorithms. II. Program Implementation of the Weisfeiler-Leman Algorithm. Technical Report TUM-M9701, TU München, Garching, Germany, 1995.

[8] R. A. Beezer. Trivalent orbit-polynomial graphs. *Lin. Alg. Appl.*, 73:133–146, 1986.

[9] J. H. Bennett. On the theory of random mating. *Ann. Eugen.*, 18:311–317, 1954.

[10] A. D. Bethke. *Genetic Algorithms and Function Optimizers.* PhD thesis, University of Michigan, 1991.

[11] R. N. Bhattacharya and E. C. Waymire. *Stochastic Processes with Applications.* Wiley, New York, 1990.

[12] N. Biggs. *Algebraic Graph Theory.* Cambridge University Press, Cambridge UK, 2nd edition, 1994.

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

[14] R. C. Bose and D. M. Mesner. On linear associative algebras corresponding to association schemes of partially balanced designs. *Ann. Math. Statist.*, 30:21–38, 1959.

[15] A. J. Bray and M. A. Moore. Metastable states in spin glasses with short-ranged interactions. *J. Phys. C*, 14:1313–1327, 1981.

[16] A. Brouwer, A. Cohen, and A. Neumaier. *Distance-regular Graphs.* Springer Verlag, Berlin, New York, 1989.

[17] F. R. K. Chung. *Spectral Graph Theory*, volume 92 of *CBMS*. American Mathematical Society, 1997.

[18] B. Codenotti and L. Margara. Local properties of some np-complete problems. Technical Report TR 92-021, International Computer Science Institute, Berkeley, CA, 1992.

[19] J. C. Culberson. Mutation-crossover isomorphism and the construction of discriminating functions. *Evol. Comp.*, 2:279–311, 1995.

[20] J. Cupal and P. F. Stadler. RNA shape space topology. in preparation, 1999.

[21] D. M. Cvetković, M. Doob, I. Gutman, and A. Torgašev. *Recent Results in the Theory of Graph Spectra*, volume 36 of *Annals of Discrete Mathematics*. North Holland, Amsterdam, New York, Oxford, Tokyo, 1988.

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

[23] D. M. Cvetković, P. Rowlinson, and S. Simić. *Eigenspaces of Graphs*, volume 66 of *Encyclopdia of Mathematics*. Cambrigdge University Press, Cambridge, UK, 1997.

[24] Y. C. de Verdière. Multiplicités des valeurs propres laplaciens discrete at laplaciens continus. *Rendiconti di Matematica*, 13:433–460, 1993.

[25] P. Delsarte. *An algebraic approach to association schemes of coding theory*, volume 10 of *Phillips Research Reports Supplements*. Phillips, 1973.

[26] B. Derrida. Random energy model: Limit of a family of disordered models. *Phys. Rev. Lett.*, 45:79–82, 1980.

[27] B. Derrida and E. Gardner. Metastable states of a spin glass chain at 0 temperature. *J. Physique*, 47:959–965, 1986.

[28] B. Derrida and L. Peliti. Evolution in a flat fitness landscape. *Bull. Math. Biol.*, 53, 1991.

[29] P. Diaconis. *Group Representations in Probability and Statistics*. Inst. of Math. Stat., Hayward, CA, 1989.

[30] W. Ebeling, A. Engel, B. Esser, and R. Feistel. Diffusion and reaction in random media and models of evolution processes. *J. Stat. Phys.*, 37:369–384, 1984.

[31] M. Eigen, J. McCaskill, and P. Schuster. The molecular Quasispecies. *Adv. Chem. Phys.*, 75:149 – 263, 1989.

[32] M. Eigen, J. McCaskill, and P. Schuster. The molecular Quasispecies. *Adv. Chem. Phys.*, 75:149 – 263, 1989.

[33] M. Eigen and P. Schuster. *The Hypercycle*. Springer-Verlag, New York, Berlin, 1979.

[34] I. A. Faradžev, M. H. Klin, and M. E. Muzychuk. Cellular rings and groups of automorphisms of graphs. In I. A. Faradžev, A. A. Ivanov, M. H. Klin, and A. J. Woldar, editors, *Investigations in Algebraic Theory of Combinatorial Objects*, volume 84 of *Mathematics and Its Applications Soviet Series*. Kluwer, Dordrecht, 1994.

[35] J. Felsenstein. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.*, 17:368–376, 1981.

[36] M. Fiedler. A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory. *Czechoslovak Math. J.*, 25:619–633, 1975.

[37] P. Field. Non-binary transformations of genetic algorithms. *Complex Systems*, 9:11–28, 1995.

[38] D. B. Fogel. *Evolutionary Computation*. IEEE Press, New York, 1995.

[39] W. Fontana, D. A. M. Konings, P. F. Stadler, and P. Schuster. Statistics of RNA secondary structures. *Biopolymers*, 33:1389–1404, 1993.

[40] W. Fontana and P. Schuster. Continuity in evolution: On the nature of transitions. *Science*, 280:1451–1455, 1998.

[41] W. Fontana and P. Schuster. Shaping space: The possible and the attainable in RNA genotype-phenotype mapping. *J. Theor. Biol.*, 1998.

[42] W. Fontana, 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 landscapes. *Phys. Rev. E*, 47:2083 – 2099, 1993.

[43] S. Forrest and M. Mitchell. Relative building block fitness and the building block hypothesis. In L. D. Whitley, editor, *Foundations of Genetic Algorithms 2*, pages 109–126. Morgan Kaufmann, San Mateo, CA, 1993.

[44] H. Frauenfelder, A. R. Bishop, A. Garcia, A. Perelson, P. Schuster, D. Sherrington, and P. J. Swart, editors. *Landscape Paradigms in Physics and Biology: Concepts, Structures, and Dynamics.* Elsevier, Amsterdam, 1997. Special Issue of *Phyica D* vol. 107(2-4).

[45] J. Friedman. Some geometric aspects of graphs and their eigenfunctions. *Duke Math. J.*, 69(3):487–525, March 1993.

[46] R. García-Pelayo and P. F. Stadler. Correlation length, isotropy, and meta-stable states. *Physica D*, 107:240–254, 1997.

[47] S. Gavrilets and J. Gravner. Percolation on the fitness hypercube and the evolution of reproductive isolation. *J. Theor. Biol.*, 184:51–64, 1997.

[48] J. H. Gillespie. *The Causes of Molecular Evolution.* Oxford University Press, New York, Oxford, 1991.

[49] P. Gitchoff and G. P. Wagner. Recombination induced hypergraphs: A new approach to mutation-recombination isomorphism. *Complexity*, 2:47–43, 1996.

[50] C. D. Godsil. *Algebraic Combinatorics.* Chapman & Hall, New York, 1993.

[51] C. D. Godsil. Equitable partitions. In D. Miklós, V. T. Sós, and T. Szőnyi, editors, *Combinatorics, Paul Erdős in Eighty*, volume 1, pages 173–192. János Bolyai Mathematical Society, Budapest, 1993.

[52] C. D. Godsil. Tools from linear algebra. In R. Graham, M. Grötschel, and L. Lovász, editors, *Handbook of Combinatorics*, pages 1105–1748. North-Holland, Amsterdam, 1995.

[53] D. E. Goldberg. Genetic algorithms and walsh functions. Part I: a gentle introduction. *Complex Systems*, 3:129–152, 1989.

[54] D. E. Goldberg. Genetic algorithms and Walsh functions. Part II: deceptiveness and its analysis. *Complex Systems*, 3:153–176, 1989.

[55] D. E. Goldberg and M. Rudnik. Genetic algorithms and the variance of fitness. *Complex Systems*, 5:265–278, 1991.

[56] D. J. Gross and M. Mèzard. The simplest spin glass. *Nucl. Phys. B*, 240:431–452, 1984.

[57] L. K. Grover. Local search and the local structure of NP-complete problems. *Oper. Res. Lett.*, 12:235–243, 1992.

[58] W. Grüner, R. Giegerich, D. Strothmann, C. M. Reidys, J. Weber, I. L. Hofacker, P. F. Stadler, and P. Schuster. Analysis of RNA sequence structure maps by exhaustive enumeration. I. Neutral networks. *Monath. Chem.*, 127:355–374, 1996.

[59] W. Grüner, R. Giegerich, D. Strothmann, C. M. Reidys, J. Weber, I. L. Hofacker, P. F. Stadler, and P. Schuster. Analysis of RNA sequence structure maps by exhaustive enumeration. II. Structures of neutral networks and shape space covering. *Monath. Chem.*, 127:375–389, 1996.

[60] R. Happel and P. F. Stadler. Canonical approximation of fitness landscapes. *Complexity*, 2:53–58, 1996.

[61] R. B. Heckendorn and D. Whitley. A Walsh analysis of NK-landscapes. In T. Baeck, editor, *International Conference on Genetic Algorithms*. Morgan Kaufmann, San Francisco CA, 1997.

[62] T. Hida. *Brownian Motion*. Springer-Verlag, New York, 1980.

[63] D. G. Higman. Intersection matrices for finite permutation groups. *J. Algebra*, 6:22–42, 1967.

[64] D. G. Higman. Coherent configurations. I: Ordinary representation theory. *Geometriae Dedicata*, 4:1–32, 1975.

[65] D. G. Higman. Coherent configurations. Part II: Weights. *Geometriae Dedicata*, 5:413–424, 1976.

[66] I. L. Hofacker, W. Fontana, P. F. Stadler, S. Bonhoeffer, M. Tacker, and P. Schuster. Fast folding and comparison of RNA secondary structures. *Monatsh. Chemie*, 125(2):167–188, 1994.

[67] A. J. Hoffman. On the polynomial of a graph. *Amer. Math. Monthly*, 70:30–36, 1963.

[68] J. H. Holland. Genetic algorithms and classifier systems: foundations and future directions. In *Proceedings of the 2nd International Conference on Genetic Algorithms*, pages 82–89, 1987.

[69] J. H. Holland. *Adaptation in Natural and Artificial Systems*. MIT Press, Cambridge, MA, 1993.

[70] W. Hordijk. A measure of landscapes. *Evolutionary Computation*, 4(4):335–360, 1996.

[71] W. Hordijk. Correlation analysis of the synchronizing-CA landscape. *Physica D*, 107:255–264, 1997.

[72] W. Hordijk and P. F. Stadler. Amplitude spectra of fitness landscapes. *J. Complex Systems*, 1:39–66, 1998.

[73] H. Hotelling. Analysis of a complex of statistical variables into principal components. *J. Educ. Psych.*, 24:417–441 and 498–520, 1933.

[74] M. A. Huynen and P. Hogeweg. Pattern generation in molecular evolution. Exploitation of the variation in RNA landscapes. *J. Mol. Evol.*, 39:71–79, 1994.

[75] M. A. Huynen, P. F. Stadler, and W. Fontana. Smoothness within ruggedness: the role of neutrality in adaptation. *Proc. Natl. Acad. Sci. (USA)*, 93:397–401, 1996.

[76] T. Jones. *Evolutionary Algorithms, Fitness Landscapes, and Search.* PhD thesis, Univ. of New Mexico, Albuquerque, NM, 1995.

[77] T. Jones. One operator, one landscape. Technical Report #95-02-025, Santa Fe Institute, 1995.

[78] S. Karlin and H. M. Taylor. *A first course in stochastic processes.* Academic Press, New York, 1975.

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

[80] S. A. Kauffman and S. Levin. Towards a general theory of adaptive walks on rugged landscapes. *J. Theor. Biol.*, 128:11–45, 1987.

[81] J. Kececioglu and D. Sankoff. Exact and approximate algorithms for sorting by reversals, with applications to genome rearrangement. *Algorithmica*, 13:180–210, 1995.

[82] W. Kern. On the depth of combinatorial optimization problems. *Discr. Appl. Math.*, 43:115–129, 1993.

[83] M. Kimura. *The Neutral Theory of Molecular Evolution.* Cambridge University Press, Cambridge, UK, 1983.

[84] G. Kirchhoff. Über die Auflösung der Gleichungen, auf welche man bei der Untersuchung der lineare Verteilung galvanischer Ströme geführt wird. *Ann. Phys. Chem.*, 72:487–508, 1847.

[85] M. Klin, C. Rücker, G. Rücker, and G. Tinhofer. Algebraic combinatorics in mathematical chemistry. Methods and algorithms. I. Permutation groups and coherent (cellular) algrebras. Technical Report TUM-M9510, TU München, Garching, Germany, 1997.

[86] M. H. Klin, A. Munemasa, M. E. Muzychuk, and P.-H. Zieschang. Directed strongly regular graphs via coherent (cellular) algebras. Ben Gurion University, preprint, 1997.

[87] M. H. Klin, R. Pöschel, and K. Rosenbaum. *Angewandte Algebra.* Vieweg, Braunschweig, 1988. german.

[88] J. R. Koza. *Genetic Programming: On the Programming of Computers by Means of Natural Selection.* MIT Press, Cambridge, MA, 1992.

[89] B. Krakhofer and P. F. Stadler. Local minima in the graph bipartitioning problem. *Europhys. Lett.*, 34:85–90, 1996.

[90] C. L.Barrett, H. Mortveit, and C. M. Reidys. Elements of a theory of simulation II: sequential dynamical systems. *Appl. Math. & Comp.*, 1998. in press.

[91] G. E. Liepins and M. D. Vose. Deceptiveness and genetic algorithm dynamics. In G. J. E. Rawlins, editor, *Foundations of Genetic Algorithms*, pages 36–50. Morgan Kaufmann, San Mateo, CA, 1991.

[92] G. E. Liepins and M. D. Vose. Polynomials, basis sets, and deceptiveness in genetic algorithms. *Complex Systems*, 5:45–61, 1991.

[93] G. E. Liepins and M. D. Vose. Characterizing crossover in genetic algorithms. *Ann. Math. Aritif. Intel.*, 5:27–34, 1992.

[94] L. Lovász. Spectra of graphs with transitive groups. *Periodica Math.Hung.*, 6:191–195, 1975.

[95] L. Lovász. Random walks on graphs: A survey. In D. Miklós, V. T. Sós, and T. Szőnyi, editors, *Combinatorics, Paul Erdős in Eighty*, volume 2, pages 1–46. János Bolyai Mathematical Society, Budapest, 1993.

[96] Y. I. Lyubich. *Mathematical structures in population genetics.* Springer-Verlag, Berlin, 1992.

[97] D. Maslen and D. Rockmore. Generalized FFTs – a survey of some recent results. In L. Finkelstein and W. Kantor, editors, *Groups and Computation II*, volume 28 of *DIMACS*, pages 183–238, Providence, RI, 1996. American Mathmatical Society.

[98] M. Mézard, G. Parisi, and M. Virasoro. *Spin Glass Theory and Beyond.* World Scientific, Singapore, 1987.

[99] B. Mohar. The Laplacian spectrum of graphs. In Y. Alavi, G. Chartrand, O. Ollermann, and A. Schwenk, editors, *Graph Theory, Combinatorics, and Applications*, pages 871–898, New York, 1991. John Wiley and Sons, Inc.

[100] B. Mohar. Some applications of Laplace eigenvalues of graphs. In G. Hahn and G. Sabidussi, editors, *Graph Symmetry: Algebraic Methods and Applications*, volume 497 of *NATO ASI Series C*, pages 227–275. Kluwer, Dordrecht, 1997.

[101] T. Nagylaki. *Introduction to Theoretical Population Genetics.* Springer-Verlag, Berlin, 1992.

[102] M. E. J. Newman and R. Engelhardt. Effects of selective neutrality on the evolution of molecular species. *Proc. Roy. Soc. (London) B*, 265:1333–1338, 1998.

[103] K. Ohmori and E. Tanaka. A unified view on tree metrics. In G. Ferrate, editor, *Syntactic and Structural Pattern Recognition*, pages 85–100, Berlin, Heidelberg, 1988. Springer-Verlag.

[104] T. Ohta. The current significance of neutral and near neutral theories. *BioEssays*, 18:673–677, 1996.

[105] U.-M. O'Reilly. Using a distance metric on genetic programs to understand genetic operators, 1997. Genetic Programming 1997, I.E.E.E. Systems, Man and Cybernetics, Orlando FL.

[106] U.-M. O'Reilly and F. Oppacher. Program search with a hierarchical variable length representation: Genetic programming, simulated annealing, and hill climbing. In Y. Davidor, H. Schwefel, and R. Manner, editors, *Parallel Problem Solving from Nature III*, volume 866 of *Lecture Notes in Computer Science*, pages 39–46. Springer, Berlin, 1994.

[107] S. E. Page and D. E. Richardson. Walsh functions, scheme variance, and deception. *Complex Systems*, 6:125–136, 1992.

[108] R. Palmer. Optimization on rugged landscapes. In A. S. Perelson and S. A. Kauffman, editors, *Molecular Evolution on Rugged Landscapes:*

Proteins, RNA, and the Immune System, pages 3–25. Addison Wesley, Redwood City, CA, 1991.

[109] J. Pearl. *Heuristics: Intelligent search strategies for computer problem solving*. Addison-Wesley, Reading MA, 1984.

[110] A. S. Perelson and S. A. Kauffman, editors. *Molecular Evolution on Rugged Landscapes: Proteins, RNA, and the Immune System*, volume 9 of *Santa Fe Institute Studies*, Reading, MA, 1991. Addison-Wesley.

[111] D. L. Powers. Eigenvectors of distance-regular graphs. *SIAM J. Matrix Anal. Appl.*, 9:399–407, 1988.

[112] D. L. Powers. Graph partitioning by eigenvectors. *Lin. Algebra Appl.*, 101:121–133, 1988.

[113] D. L. Powers and M. M. Sulaiman. The walk partition and colorations of a graph. *Linear Algebra Appl.*, 48:145–159, 1982.

[114] W. B. Provine. *Sewall Wright and Evolutionary Biology*. Univ. Chicago Press, Chicago, London, 1986.

[115] C. M. Reidys. Random induced subgraphs of generalized n-cubes. *Adv. Appl. Math.*, 19:360–377, 1997.

[116] C. M. Reidys. Random-structures. *Annals Comb.*, 1998. in press.

[117] C. M. Reidys, C. V. Forst, and P. Schuster. Replication and mutation on neutral networks of RNA secondary structures. *Bull. Math. Biol.*, 1998. in press, Santa Fe Institute Preprint 98-04-036.

[118] C. M. Reidys and P. F. Stadler. Neutrality in fitness landscapes. *Appl. Math. & Comput.*, 1998. submitted, Santa Fe Institute preprint 98-10-089.

[119] C. M. Reidys, P. F. Stadler, and P. Schuster. Generic properties of combinatory maps: Neural networks of RNA secondary structures. *Bull. Math. Biol.*, 59:339–397, 1997.

[120] H. Rieger. The number of solutions of the Thouless-Anderson-Palmer equations for p-spin interaction spin glasses. *Phys. Rev. B*, 46:14655–14661, 1992.

[121] R. B. Robbins. Some applications of mathematics to breeding problems. III. *Genetics*, 3:375–389, 1918.

[122] D. Rockmore. Some applications of generalized FFTs. In L. Finkelstein and W. Kantor, editors, *Groups and Computation II*, volume 28 of *DIMACS*, pages 329–370, Providence, RI, 1995. American Mathmatical Society.

[123] F. Runge and H. Sachs. Berechnung der Anzahl der Gerüste von Graphen und Hypergraphen mittels deren Spektren. *Math. Balkanica (Belgrade)*, 4:529–536, 1974.

[124] J. Ryan. The depth and width of local minima in discrete solution spaces. *Discr. Appl. Math.*, 56:75–82, 1995.

[125] P. Schuster, W. Fontana, P. F. Stadler, and I. L. Hofacker. From sequences to shapes and back: A case study in RNA secondary structures. *Proc. Roy. Soc. Lond. B*, 255:279–284, 1994.

[126] P. Schuster and P. F. Stadler. Landscapes: Complex optimization problems and biopolymer structures. *Computers & Chem.*, 18:295–314, 1994.

[127] P. Schuster, P. F. Stadler, and A. Renner. RNA structures and folding: From conventional to new issues in structure predictions. *Curr. Opinions Structural Biol.*, 7:229–235, 1997.

[128] A. Schwenk. Computing the characteristic polynomial of a graph. In *Graphs and Combinatorics*, volume 406 of *Lecture Notes in Mathematics*, pages 153–162. Springer-Verlag, Berlin, 1974.

[129] D. Sherrington and S. Kirkpatrick. Solvable model of a spin-glass. *Physical Review Letters*, 35(26):1792 – 1795, 1975.

[130] M. Shpak and G. P. Wagner. Asymmetry of configuration spaces induced by unequal crossover: Implications for the mathematical theory of evolutionary innovation. Preprint, Dept. of Biology, Yale University, New Haven CT, 1999.

[131] G. B. Sorkin. Combinatorial optimization, simulated annealing, and fractals. Technical Report RC13674 (No.61253), IBM Research Report, 1988.

[132] P. F. Stadler. Correlation in landscapes of combinatorial optimization problems. *Europhys. Lett.*, 20:479–482, 1992.

[133] P. F. Stadler. Linear operators on correlated landscapes. *J. Physique I (France)*, 4:681–696, 1994.

[134] P. F. Stadler. Random walks and orthogonal functions associated with highly symmetric graphs. *Discr. Math.*, 145:229–238, 1995.

[135] P. F. Stadler. Towards a theory of landscapes. In R. Lopéz-Peña, R. Capovilla, R. García-Pelayo, H. Waelbroeck, and F. Zertuche, editors, *Complex Systems and Binary Networks*, pages 77–163, Berlin, New York, 1995. Springer Verlag.

[136] P. F. Stadler. Landscapes and their correlation functions. *J. Math. Chem.*, 20:1–45, 1996.

[137] P. F. Stadler. Fitness landscapes arising from the sequence-structure maps of biopolymers. *J. Mol. Struct. (THEOCHEM)*, 1998. in press, Sante Fe Institute Preprint 97-11-082.

[138] P. F. Stadler and R. Happel. Correlation structure of the landscape of the graph-bipartitioning-problem. *J. Phys. A: Math. Gen.*, 25:3103–3110, 1992.

[139] P. F. Stadler and R. Happel. Random field models for fitness landscapes. *J. Math. Biol.*, 1999. in press, SFI preprint 95-07-069.

[140] P. F. Stadler and B. Krakhofer. Local minima of p-spin models. *Rev. Mex. Fis.*, 42:355–363, 1996.

[141] P. F. Stadler and W. Schnabl. The landscape of the traveling salesman problem. *Phys. Letters A*, 161:337–344, 1992.

[142] P. F. Stadler, R. Seitz, and G. P. Wagner. Evolvability of complex characters: Population dependent fourier decomposition of fitness landscapes over recombination spaces. *Bull. Math. Biol.*, 1999. submitted, Santa Fe Institute Preprint 99-01-001.

[143] P. F. Stadler and G. P. Wagner. The algebraic theory of recombination spaces. *Evol. Comp.*, 5:241–275, 1998.
[144] C. R. Stephens and H. Waelbroeck. Effective degrees of freedom in genetic algorithms. *Phys. Rev. E*, 57:3251–3264, 1998.
[145] M. Tacker, P. F. Stadler, E. G. Bornberg-Bauer, I. L. Hofacker, and P. Schuster. Algorithm independent properties of RNA structure prediction. *Eur. Biophy. J.*, 25:115–130, 1996.
[146] K. Tai. The tree-to-tree correction problem. *J. ACM*, 26:422–433, 1979.
[147] F. Tanaka and S. F. Edwards. Analytic theory of ground state properties of a spin glass: I. Ising spin glass. *J. Phys. F*, 10:2769–2778, 1980.
[148] D. J. Thouless, P. W. Anderson, and R. G. Palmer. Solution of 'Solvable model of a spin glass'. *Phil. Mag.*, 35:593–601, 1977.
[149] H. van der Holst. *Topological and Spectral Graph Characterizations.* PhD thesis, Universiteit van Amsterdam, 1996.
[150] E. van Nimwegen and J. P. Crutchfield. Optimizing epochal evolutionary search: Population-size independent theory. *Computer Methods in Applied Mechanics and Engineering*, 1999. to appear.
[151] M. D. Vose and A. H. Wright. The simple genetic algorithm and the Walsh transform. Part I: Theory. *Evol. Comp.*, 6:253–274, 1998.
[152] M. D. Vose and A. H. Wright. The simple genetic algorithm and the Walsh transform. Part II: The inverse. *Evol. Comp.*, 6:275–289, 1998.
[153] G. P. Wagner and P. F. Stadler. Complex adaptations and the structure of recombination spaces. In C. Nehaniv and M. Ito, editors, *Algebraic Engineering*, Singapore, 1998. World Scientific. Proceedings of the Conference on Semi-Groups and Algebraic Engineering, University of Aizu, Japan, in press, Santa Fe Institute Preprint 97-03-029.
[154] E. D. Weinberger. Correlated and uncorrelated fitness landscapes and how to tell the difference. *Biol. Cybern.*, 63:325–336, 1990.
[155] E. D. Weinberger. Fourier and Taylor series on fitness landscapes. *Biological Cybernetics*, 65:321–330, 1991.
[156] E. D. Weinberger and P. F. Stadler. Why *some* fitness landscapes are fractal. *J. Theor. Biol.*, 163:255–275, 1993.
[157] B. Y. Weisfeiler and A. A. Leman. Reduction of a graph to a canonical form and an algebra arising during this reduction. *Naucho – Techn. Inf.; Ser. 2*, 9:12–16, 1968. Russian.
[158] L. D. Whitley. Fundamental principles of deception in genetic search. In G. Rawlins, editor, *Foundations of Genetic Algorithms*, pages 221–241. Morgan Kaufmann, San Mateo, CA, 1991.
[159] L. D. Whitley and J. Dzubera. Advance correlation analysis of operators for the t raveling salesman problems. In Y. Davidor, H. P. Schwefel, and R. Manner, editors, *Parallel Problem Solving from Nature – PPSN III*, pages 68–77. Springer-Verlag, 1994.
[160] H. Wielandt. *Finite Permutation Groups.* Academic Press, New York, 1964.

[161] S. W. Wilson. GA-easy does not imply steepest-ascent optimizable. In R. K. Belew and L. B. Booker, editors, *Proceedings of the Fourth International Conference on Genetic Algorithms*, pages 85–89. Morgan Kaufmann, San Mateo, CA, 1991.

[162] D. H. Wolpert and W. G. Macready. No free lunch theorems for optimization. *IEEE Trans. Evol. Comp.*, 1:67–83, 1997.

[163] F. Yates. The design and analysis of factorial experiments. *Imp. Bur. Soil Sci. Tech. Comm.*, 35, 1937.