Landscapes and Their Correlation Functions

Peter F. Stadler[a,b,*]

[a]Theoretische Biochemie, Institut für Theoretische Chemie
Universität Wien, Vienna, Austria

[b]Santa Fe Institute, Santa Fe, NM

[*]Mailing Address:
Peter F Stadler, Inst. f. Theoretische Chemie, Univ. Wien
Währingerstr. 17, A-1090 Wien, Austria
Phone: [43] 1 40480 665 Fax: [43] 1 40480 660
E-Mail: `studla@tbi.univie.ac.at`

Keywords

Coherent Configuration — Combinatorial Optimization — Correlation Function — Fourier Series — Landscapes — Laplace Operator — Random Walk

Abstract

Fitness landscapes are an important concept in molecular evolution. Many important examples of landscapes in physics and combinatorial optimization, which are widely used as model landscapes in simulations of molecular evolution and adaptation, are "elementary", i.e., they are (up to an additive constant) eigenfunctions of a graph Laplacian. It is shown that elementary landscapes are characterized by their correlation functions. The correlation functions are in turn uniquely determined by the geometry of the underlying configuration space and the nearest neighbor correlation of the elementary landscape. Two types of correlation functions are investigated here: the correlation of a time series sampled along a random walk on the landscape and the correlation function with respect to a partition of the set of all vertex pairs.

1. Introduction

Since Sewall Wright's seminal paper [1] the notion of a *fitness landscape* underlying the dynamics of evolutionary optimization has proved to be one of the most powerful concepts in evolutionary theory. Implicit in this idea is a collection of genotypes arranged in an abstract metric space, with each genotype next to those other genotypes which can be reached by a single mutation, as well as a value assigned to each genotype. Such a construction is by no means restricted to biological evolution; Hamiltonians of disordered systems, such as spin glasses [2, 3], and the cost functions of combinatorial optimization problems [4] have the same basic structure. It has been known since Eigen's [5] pioneering work on the molecular quasispecies that the dynamics of evolutionary adaptation (optimization) on a landscape depends crucially on detailed structure of the landscapes itself. Extensive computer simulations, see, e.g., [6, 7] have made it very clear that a complete understanding of the dynamics is impossible without a thorough investigation of the underlying landscape [8].

The landscapes of a number of well known combinatorial optimization problems such as the Traveling Salesman Problem `TSP` [9], the Graph Bipartitioning Problem `GBP` [10], or the Graph Matching Problem `GMP` have been investigated in some detail, see [11, 12, 13]. A detailed survey of a variety of model landscapes derived from folding RNA molecules into their secondary structures has been performed recently [14, 6, 7, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24].

Most of the knowledge about landscapes has so far been derived using statistical methods, considering random models of landscapes rather than a single landscape. The distribution of local optima and the statistical characteristics of down-hill walks have been computed for the uncorrelated landscape of the random energy model [25, 26, 27], Furthermore, two one-parameter families of tunably rugged landscapes have been studied extensively: the Nk model and its variants [28, 29, 30, 16] and the p-spin models [31, 32, 33, 34]. Local optima of 2-spin models are considered in [35, 36, 37]. While the statistical approach is the natural one, e.g., in the physics of spin glasses, it seems to be rather contrived in evolutionary biology because it is by no means clear what a reasonable statistical model should look like, even if there is a computational procedure to model *the* landscapes of, say, RNA free energies.

A theory of landscapes is based on three ingredients: we are given a *finite*, but very large set V of "configurations" and a "fitness function" $f : V \to \mathbb{R}$. The third ingredient is a notion of neighborhood between the configurations, which allows us to interpret V as the vertex set of a graph Γ. We will refer to Γ as the *configuration space* of the landscape f. Let us briefly discuss two examples here:

Consider the set of RNA molecules of given chain length n. A particular molecule x can be represented as a string of length n taken from the alphabet $\{$**G**, **C**, **A**, **U**$\}$; molecular biologists call this string the *sequence* of the RNA. The "fitness" function f is, for instance, the free energy of folding x into its secondary structure [15]. *In silico* the folding is done by an algorithm containing a large number of experimentally determined parameters [38]. In nature as well as in *in vitro* experiments variation is introduced by mutations, predominantly point-mutations. Neighboring sequences are thus those that differ in only a single position. The resulting graph is known as the sequence space [39, 40].

A very different example is the Travelling Salesman Problem. A salesman starts from his home city and visits exactly once each of the n cities on a given list, then he returns home. The configurations are the possible tours, i.e., all permutations of the cities on the salesman's list. The numerical value assigned to a particular tour

τ is its total length $f(\tau)$. The notion of neighborhood between different tours is much less obvious here than in the biological example above. Usually one says that two tours are neighbors if they can be interconverted by a simple operation on the list of cities, such as swapping two cities (transpositions), or inverting the order of a contiguous part of the list. It turns out that the performance of an optimization heuristic depends crucially on the choice of the neighborhood relation. We will return to this topic later in this contribution.

The main theme of this paper is the relationship between the Fourier-expansion of a given landscapes f and its autocorrelation functions. Landscapes that are (up to an additive constant) eigenvectors of the Laplacian operator of the configuration space Γ take center stage in the theory presented here. We shall refer to them as *elementary landscapes*. Two types of correlation functions will be considered here: The autocorrelation function $r(s)$ of a "time series" generated by a random walk on the configuration space, and the correlation function $\rho(\mathcal{X})$ defined on suitable partitions of the set of all pairs of configurations. We will show that a landscape is elementary if and only if $r(s)$ is exponential which is in turn equivalent to $\rho(\mathcal{X})$ being a left eigenvector of what is called the "collapsed adjacency matrix" of the configuration space.

In section 2 we consider the properties of $r(s)$ in some detail. Section 3 is devoted to the Laplacian operator on graphs and the Fourier expansion of landscapes. We shall also prove the first part of the main theorem at this stage. Section 4 discusses the known elementary landscapes and reviews some the implications of elementarity. Correlation measures defined on partitions of the set of pairs of configurations are the subject of section 5. The second part of the main theorem will be proved there. The results of this contribution are summarized in section 6 and discussed in section 7.

2. "Random Walk" Correlation Functions of Landscapes

Because of the extremely large number of configurations, 4^n for RNA and $n!$ for the TSP for example, we need a condensed description of a landscape. Correlation measures relating the values of nearby configurations with each other seem to be a natural approach. We will show that the most useful definition of such a measure depends on the symmetry properties of the graph $\Gamma = (V, E)$, i.e., on the choice of the neighborhood relations.

Definition. For each landscape $f : V \to \mathbb{R}$ we define

$$\overline{f} \stackrel{\text{def}}{=} \frac{1}{|V|} \sum_{x \in V} f(x) \qquad \text{and} \qquad \sigma_f^2 \stackrel{\text{def}}{=} \frac{1}{|V|} \sum_{x \in V} \left[f(x) - \overline{f} \right]^2 = \overline{f^2} - \overline{f}^2 .$$

A landscape with $\sigma_f^2 = 0$ is called *flat.*

Note that $\sigma_f^2 = 0$ if and only if f is constant, i.e., for *flat* landscapes. The quantity $\overline{f}$ is the mean of the landscape, and σ_f^2 can be interpreted as the variance of the landscape. There is nothing "statistical" about $\overline{f}$ or σ_f^2. Both quantities are well defined functionals of $f : V \to \mathbb{R}$, and they should not be mistaken for the averages over different instances which are commonly used in the analysis of statistical models of landscapes.

Both $\overline{f}$ and σ_f^2 do not depend on the neighborhood structure implied by the edge set E of Γ. Eigen and co-workers [8] have introduced correlation functions that depend on the Hamming distance in sequence space as a measure of the local structure of fitness landscapes. Weinberger [41] proposed to use a simple random walk $\{x_0, x_1, \ldots\}$ on the vertex set in order to sample a "time series" $\{f(x_0), f(x_1), \ldots\}$ and to use the autocorrelation function of this "time series" as a characteristic of the landscape. A simple random walk [42] on a graph Γ has transition matrix $\mathbf{T} \stackrel{\text{def}}{=} \mathbf{D}^{-1}\mathbf{A}$, where $\mathbf{A}$ is the adjacency matrix of Γ and $\mathbf{D}$ is the diagonal matrix of vertex degrees with entries

$$\mathbf{A}_{xy} \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } \{x,y\} \in E \\ 0 & \text{otherwise} \end{cases} \qquad \text{and} \quad \mathbf{D}_{xy} \stackrel{\text{def}}{=} \delta_{x,y} \sum_{z \in V} \mathbf{A}_{xz} = \delta_{x,y} [\mathbf{A1}]_x \, ,$$

respectively. Here $\delta_{x,y}$ is the Kronecker symbol, and $\mathbf{1} \stackrel{\text{def}}{=} (1,1,\dots,1)$ is the vector with all entries 1.

The expected autocorrelation function of a "time series" sampled along a simple random walk on Γ is defined as

$$r(s) \stackrel{\text{def}}{=} \frac{\langle f(x_t)f(x_{t+s})\rangle_{x_0,t} - \langle f(x_t)\rangle_{x_0,t}\,\langle f(x_{t+s})\rangle_{x_0,t}}{\sqrt{\left(\langle f(x_t)^2\rangle_{x_0,t} - \langle f(x_t)\rangle^2_{x_0,t}\right)\left(\langle f(x_{t+s})^2\rangle_{x_0,t} - \langle f(x_{t+s})\rangle^2_{x_0,t}\right)}}$$

where the notation $\langle\,.\,\rangle_{x_0,t}$ emphasizes that the expectation is taken over all "times" t and all initial conditions x_0. We will refer to $r(s)$ as the *"random walk" correlation function* of the landscape f on Γ. Since the averages are taken over all initial conditions with uniform weights, the definition of $r(s)$ simplifies to

$$r(s) = \frac{\langle f(x_t)f(x_{t+s})\rangle_{x_0,t} - \langle f(x_t)\rangle^2_{x_0,t}}{\langle f(x_t)^2\rangle_{x_0,t} - \langle f(x_t)\rangle^2_{x_0,t}} \tag{1}$$

Note that, again, $r(s)$ can be viewed as a functional of $f: V \to \mathbb{R}$. While the definition of $r(s)$ via the random walk is convenient for sampling data from a given landscape, say a particular instance of a combinatorial optimization problem or a model of the complicated energy function for protein folding, it is inconvenient from a mathematical point of view. Furthermore, it seems to be rather contrived to invoke a stochastic process in order to characterize a given function defined on a finite set. We will therefore derive a representation of r that does not require the explicit assumption of a random walk, at least for the case of regular graphs. It is not at all clear that $r(s)$ is a useful measure of a landscape on a more irregular graph. To this end we need two preparatory lemmata[1].

Lemma 1. Let Γ be a regular graph and let $F: V \to \mathbb{R}$ be an arbitrary function. Let $\{x_t\}$ be a simple random walk on Γ. Then $\langle F(x_t)\rangle_{x_0,t} = \overline{F}$.

[1]In this contribution we will label a technical result as a "proposition" if it is known in the literature, and as a "lemma" otherwise. The proofs of all lemmata and their corollaries can be found in the appendix.

Lemma 2. Let Γ be a regular graph, and let $F : V \times V \to \mathbb{R}$. Then

$$\langle F(x_{t+s}, x_t) \rangle_{x_0,t} = \frac{1}{|V|} \sum_{x,y \in V} F(x,y) [\mathbf{T}]^s_{xy} \,.$$

The desired representation of $r(s)$ is now obtained as a corollary of the above two lemmata.

Corollary 1. Let $f : V \to \mathbb{R}$ be a non-flat landscape on a D-regular graph Γ with adjacency matrix $\mathbf{A}$. Then

$$r(s) = \frac{\frac{1}{|V|}\langle f, \mathbf{T} f\rangle - \overline{f}^2}{\overline{f^2} - \overline{f}^2} \tag{2}$$

where $\mathbf{T} = (1/D)\mathbf{A}$.

Remark. Both for the proof of lemma 1 and the proof of lemma 2 it is crucial that $\mathbf{T1} = \mathbf{1}$. Since each transition necessarily ends in some vertex $x \in V$ we have always $\mathbf{1T} = \mathbf{1}$. In other words, instead of insisting the Γ be a regular graph and that the random walk is simple we could as well require that $\mathbf{T}$ is a bi-stochastic, but not necessarily symmetric, transition matrix. This might be a starting point for an investigation of landscapes on non-regular configuration spaces.

3. Graph Laplacians and Their Eigenfunctions

3.1. Incidence Matrix and Graph Laplacian

The incidence matrix and the Laplacian matrix of a graph can be viewed as discrete analogues of the gradient and the Laplace operator in Euclidean spaces.

Definition. Let Γ be an arbitrary graph with vertex set V and edge set E. For each edge $h = \{v, w\}$ we choose one of the two vertices as the "positive end" and the other one as the "negative end" of the edge. The choice of this orientation is completely arbitrary. The matrix ∇^+ with entries

$$\nabla^+_{ij} = \begin{cases} +1 & \text{vertex } v_i \quad \text{is the positive end of edge } e_j \\ -1 & \text{vertex } v_i \quad \text{is the negative end of edge } e_j \\ 0 & \text{otherwise} \end{cases}$$

is called the *incidence matrix* of Γ.

The choice of the symbol ∇ is intentional. In fact, let $f : V \to \mathbb{R}$ be an arbitrary landscape. Then $\nabla f : E \to \mathbb{R}$ is given by

$$(\nabla f)(h) = f(v) - f(w) \qquad \text{where } h = \{v, w\}$$

and v is the positive end of the edge h. This is as close to a differential operator as one can get on a graph.

Definition. Let $\mathbf{D}$ be the diagonal matrix of vertex degrees, i.e., $\mathbf{D}_{xx}$ is the number of edges incident into x, and let $\mathbf{A}$ be the adjacency matrix of Γ. Then the matrix

$$-\Delta = \mathbf{D} - \mathbf{A} \tag{3}$$

is called the *Laplacian* of Γ. For D-regular graphs, i.e., graphs for which all vertices have degree D, this becomes $\Delta = \mathbf{A} - D\mathbf{I}$.

The graph Laplacian shares its most important properties with the familiar differential operator $\Delta = \sum_{i=1}^n \frac{\partial^2}{\partial x_i^2}$ in $\mathbb{R}^n$, as explained in some more detail in figure 1.

Proposition 1.

(i) Δ is symmetric.

(ii) $-\Delta$ is non-negative definite.

(iii) Δ is singular; the eigenvector $\mathbf{1} = (1, \dots, 1)$ belongs to the eigenvalue $\lambda_0 = 0$. If Γ is connected (as we will always assume), then λ_0 has multiplicity 1.

(iv) $-\Delta = \nabla^+\nabla$, that is, it corresponds to "second derivatives" on the graphs.

(v) For any two landscapes f and g *Green's formula* holds in the following form

$$\sum_{x\in V} f(x)(\Delta g)(x) = \sum_{x\in V} g(x)(\Delta f)(x) = -\sum_{h\in E} (\nabla f)(h)(\nabla g)(h).$$

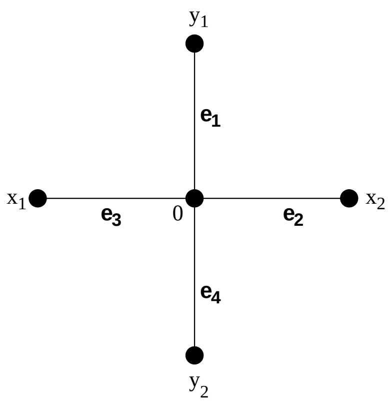

Figure 1: The graph Laplacian Δ is a generalization of the discrete approximation of the familiar Laplacian differential operator in $\mathbb{R}^n$.
This approximation is commonly performed by replacing the continuous space by a square lattice. Rescaling the space coordinates we can assume that the lattices points have integer coordinates. The first derivatives, evaluated at the mid-points of the edges, e_i, are computed from

$$\frac{\partial f}{\partial x}(e_2) = f(x_2) - f(0) \qquad \frac{\partial f}{\partial x}(e_3) = f(0) - f(x_1),$$

and analogous expressions for $\partial f/\partial y$. Consequently, the second derivative evaluated in 0 become

$$\frac{\partial^2 f}{\partial x^2}(0) = (f(x_2) - f(0)) - (f(0) - f(x_1)) = f(x_1) + f(x_2) - 2f(0),$$

and an analogous expression for $\partial^2 f/\partial y^2$. The discrete approximation of the usual Laplacian hence coincides with the graph Laplacian of the square lattice.

Proof. (i) is obvious, (ii) and (iii) are well known, see, e.g., [43, 44]. Claim (iv) is Proposition 4.8 of [43]. Green's formula (v) is easily checked by explicit calculation:

$$\begin{aligned}\sum_{h\in E}(\nabla f)(h)\,(\nabla g)(h) &= \sum_{x\in V}\sum_{y\in V}\sum_{h\in E} f(x)\nabla_{hx}\nabla_{hy}g(y)\\ &= \sum_{x\in V}\sum_{y\in V} f(x)\left(\sum_{h\in E}\nabla^+_{xh}\nabla_{hy}\right)g(y)\\ &= -\sum_{x\in V}\sum_{y\in V} f(x)\Delta_{xy}g(y) = -\sum_{x\in V} f(x)(\Delta g)(x),\end{aligned}$$

A similar calculation shows $\sum_h(\nabla f)(h)(\nabla g)(h) = \sum_x g(x)(\Delta f)(x)$. ∎

The graph Laplacian is central to the theory of electrical networks. As a reference we give Kirchhoff's classical paper [45]. Let $\eta(x)$ be the current flowing into the network at vertex x. Then there is a potential $\Phi : V \to \mathbb{R}$ satisfying $\Delta\Phi = \eta$, and the vector ζ of currents along the edges is given by $\zeta = \nabla\Phi$. A recent book on potential theory on discrete lattices is [46].

Finally we note the following connection between the spectrum of a graph and the graph Laplacian: Suppose Γ is D-regular with adjacency matrix $\mathbf{A}$. Then $\mathbf{A}$ and $-\Delta$ have the same eigenvectors, and thus the eigenvectors of $-\Delta$ are given by $-\lambda_k = \mu_k - D$.

3.2. Fourier Expansion of Landscapes

A series expansion in terms of a complete and orthonormal system of eigenfunctions of the Laplace operator is commonly termed Fourier expansion. We will adopt the same terminology here following [30]. Thus, let f be a landscape on Γ and let $\{\varphi_i\}$ denote a complete orthonormal set of eigenvectors of the graph Laplacian $-\Delta$. Then we call the expansion

$$f(x) = \sum_{i=1}^{|V|} a_i \varphi_i(x) \tag{4}$$

a *Fourier expansion* of the landscape. It will often be convenient to label the eigenvectors φ_i by the vertices of the underlying graph Γ. This is possible because the eigensystem of the finite symmetric matrix Δ is complete. In general, this labeling is of course arbitrary.

Since we deal with a finite vector space with a scalar product, for which we will use the notation $\langle\,.\,,\,.\,\rangle$, spanned by eigenvectors $\{\varphi_i\}$ of the graph Laplacian, the familiar properties of Fourier series, such as Parseval's equation

$$\|f\|^2 = \langle f, f\rangle = \sum_{y\in V} \left(\frac{\langle f, \varphi_y\rangle}{\langle \varphi_y, \varphi_y\rangle}\right)^2 \;.$$

Another important result is the mean square approximation theorem:

Proposition 2. Consider a landscape f on Γ with Fourier expansion $f = \sum_{y\in V} a_y\varphi_y$. Let X be a subset of V, and consider approximations of f of the form $g = \sum_{y\in X} b_y\varphi_y$. Then the squared approximation error $\|f-g\|^2 = \langle (f-g),(f-g)\rangle$ is minimized by choosing $b_y = a_y \equiv \langle f,\varphi_y\rangle$ for all $y \in X$.

It is clear that landscapes which are eigenfunctions of the graph Laplacian will play a special role. It will turn out to be more useful, however, to consider a slightly larger class of landscapes.

Definition. A landscape $f : V \to \mathbb{R}$ is *elementary* if there are constants f^* and λ such that

$$\Delta f + \lambda(f - f^*\mathbf{1}) = 0 \tag{5}$$

This definition is motivated by Lov Grover's observation [47] that the cost functions of a number of well-studied combinatorial optimization problems in fact fulfill equ.(5). This will be discussed in detail in section 4.

Lemma 3. A non-flat landscape on a connected graph Γ is elementary if and only if

$$f(x) = \overline{f} + \varphi(x) \qquad \forall x \in V\,, \tag{6}$$

where φ is an eigenfunction of $-\Delta$ with eigenvalue $\lambda > 0$.

3.3. "Random Walk" Correlation Functions of Elementary Landscapes

The "random walk" correlation function $r(s)$ provides an elegant way of characterizing elementary landscapes.

Theorem 1. Let f be a non-flat landscape on a D-regular graph Γ and let $r(s)$ be the "random walk" correlation function of f. Then f is elementary if and only if $r(s)$ is an exponential function, i.e., iff $r(s) = \varrho^s$.

Proof. Let $\{\varphi\}$ be an orthonormal set of real eigenvectors of the Laplacian on Γ, i.e., $(-\Delta)\varphi_i = \lambda_i\varphi_i$. Since

$$\mathbf{T} = \mathbf{I} + \mathbf{D}^{-1}\Delta = \mathbf{I} + \frac{1}{D}\Delta \tag{7}$$

we have $\mathbf{T}\varphi_i = 1 - \lambda_i/D$. Now substitute the decomposition $f = \sum_i a_i\varphi_i$ into the definition of $r(s)$. One finds

$$r(s) = \frac{\frac{1}{|V|}\sum_{i,j} a_i a_j \langle\varphi_i,\varphi_j\rangle(1-\lambda_i/D)^s - \left(\sum_i a_i \frac{1}{|V|}\sum_{x\in V}\varphi_i(x)\right)^2}{\frac{1}{|V|}\sum_{i,j} a_i a_j \sum_{x\in V}\varphi_i(x)\varphi_j(x) - \left(\sum_i a_i \frac{1}{|V|}\sum_{x\in V}\varphi_i(x)\right)^2}$$

Recall that $\mathbf{1}$ is always an eigenvector of Δ, belonging to $\lambda_0 = 0$. By orthogonality we have therefore $\sum_{x\in V}\varphi_i(x) = 0$ for all $i \neq 0$, and consequently $\overline{f} = a_0\overline{\varphi_0}$. Noting that $\sum_{x\in V}\varphi_i(x)\varphi_j(x) = \langle\varphi_i,\varphi_j\rangle = \delta_{ij}$ we obtain

$$r(s) = \frac{\frac{1}{|V|}\sum_i |a_i|^2(1-\lambda_i/D)^s - |a_0|^2\overline{\varphi_0}^2}{\frac{1}{|V|}\sum_i |a_i|^2 - |a_0|^2\overline{\varphi_0}^2}$$

It remains to compute $\overline{\varphi_0}$. We know $\varphi_0 = c\mathbf{1}$ with some constant $c \neq 0$. Normalization implies $1 = c^2\langle\mathbf{1},\mathbf{1}\rangle = c^2|V|$, i.e, $\varphi_0(x) = \frac{1}{\sqrt{|V|}}$ for all $x \in V$. Substituting this into $r(s)$ yields

$$r(s) = \frac{\frac{1}{|V|}|a_0|^2 + \frac{1}{|V|}\sum_{i\neq 0}|a_i|^2(1-\lambda_i/D)^s - |a_0|^2\frac{1}{|V|}}{\frac{1}{|V|}|a_0|^2 + \frac{1}{|V|}\sum_{i\neq 0}|a_i|^2 - |a_0|^2\frac{1}{|V|}}$$

It is convenient to introduce the normalized amplitudes

$$A_i \stackrel{\text{def}}{=} \frac{|a_i|^2}{\sum_{j\neq 0}|a_j|^2}, . \tag{8}$$

Note that a landscape is flat if and only if $A_i = 0$ for all $i \neq 0$. Thus the normalized amplitudes are in fact well defined for all non-flat landscapes. Furthermore $A_i = 0$ is true for all $i \neq 0$ if and only if $a_i = 0$. The expression for $r(s)$ simplifies considerably:

$$r(s) = \sum_{i\neq 0} A_i(1-\lambda_i/D)^s\,. \tag{9}$$

Consequently $r(s)$ is an exponential function if and only if all nonzero A_i belong to a single eigenvalue λ_k of Δ. This is the case if and only if f is of the form $f = (a_0/\sqrt{|V|})\mathbf{1}+\varphi$ where φ is an eigenvector of $-\Delta$. Applying lemma 3 completes the proof. ∎

The "random walk" correlation function of an elementary landscape is determined by the single parameter $\varrho \stackrel{\text{def}}{=} r(1)$, which one might call the *nearest-neighbor correlation of the landscape.* We have $\varrho = (1-\lambda_k/D)$, where λ_k is a non-zero eigenvalue of the graph Laplacian $-\Delta$. Since $r(s)$ is exponential we can define a *correlation length* ℓ by

$$\ell \stackrel{\text{def}}{=} \begin{cases} 0 & \text{if} \quad \varrho = 0 \\ -\dfrac{1}{\ln|\varrho|} & \text{if} \quad \varrho \neq 0 \end{cases} \tag{10}$$

Thus the "random walk" correlation function is of the form $r(s) = \exp(-s/\ell)$ for $\varrho > 0$, $r(s) = (-1)^s \exp(-s/\ell)$ for $\varrho < 0$, and $r(s) = \delta_{s,0}$ for $\varrho = 0$. Table 1 at the end of the following section compiles numerical values of ϱ and ℓ for a few landscapes of practical interest.

4. Elementary Landscapes

In this section we will briefly discuss a number of landscapes, all of which are elementary (or at least almost elementary). We will conveniently subdivide our discussion according to the type of the configuration space underlying the landscapes. Three classes of configuration spaces are of particular importance: (a) landscapes defined on sequences, (b) landscapes defined on permutations, and (c) landscapes defined on a set of subsets of given finite set. In this contribution we typeset matrices only boldface if they are related to the configuration space Γ in some way. Many of the landscape models discussed in the following contain matrices of parameters which will be typeset in italics.

4.1. Cayley Graphs and Cartesian Products of Graphs

An important class of graphs with high symmetry are obtained from finite groups.

Definition. Let $(G, \circ)$ be a finite group, and let Φ be a set of generators[2] of G such that (i) the group identity $\imath$ is not contained in Φ, and (ii) for each $x \in \Phi$ the inverse group element x^{-1} is also contained in Φ. The Cayley graph $\Gamma(G, \Phi)$ is the graph with vertex set G and $\{x, y\} \in E$ if and only if there is a $g \in \Phi$ such that $y = gx$, i.e., if and only if $xy^{-1} \in \Phi$.

The set of generators Φ can be interpreted as the set of all possible elementary mutations, or — in the context of an optimization heuristic — as the move set of the algorithm.

Many, but by no means all, of the configuration spaces encountered in this contribution are Cayley graphs, in particular the Hamming graphs (generalized hypercubes) and the Cayley graphs of the symmetric group S_n. We will return to their graph-theoretic properties later in this contribution. In this section we will be content with showing that a variety of interesting landscapes are elementary on appropriate graphs.

For Cayley graphs of commutative groups the eigenvectors and eigenvalues have a particularly simple form. We exploit the fact that a finite commutative group has a unique decomposition into cyclic groups, see e.g. [48, §13]: Let N_k be the orders of the cyclic groups. The cyclic group of order N_k in turn are isomorphic to the additive group modulo N_k, and thus G is isomorphic to the group of "vectors" $x = (x_1, x_2, \dots, x_m)$, $0 \le x_k < N_k$, under component-wise additions modulo N_k:

$$x \circ y = (x_1 + y_1 \bmod N_1, x_2 + y_2 \bmod N_2, \dots, x_m + y_m \bmod N_m).$$

The characters [49] of the commutative group G are given by

$$\chi_g(x) = \exp\left(2\pi i \sum_k \frac{x_k g_k}{N_k}\right) . \tag{11}$$

[2] $\Phi \subset G$ is a set of generators if each group element $z \in G$ can be represented as a finite product of elements of Φ.

Let $\Gamma(G,\Phi)$ a Cayley graph of G with graph Laplacian $-\Delta$. It is convenient to allow also for complex eigenvectors of the symmetric matrix $-\Delta$, since it can be shown [50] that

$$(-\Delta)\chi_g = \lambda_g \chi_g \qquad \text{holds with} \quad \lambda_g = \sum_{x\in\Phi} [1-\chi_g(x)] \tag{12}$$

In other words, the characters χ_g of G are the eigenvectors of any Cayley graph derived from the commutative group G.

Probably the simplest examples of Cayley graphs with commutative groups are the *complete graphs* K_n. Let G be a commutative group with n elements, for instance a cyclic group, and define $\Phi = G \setminus \{\imath\}$, where $\imath$ is as usual the group identity. Then $\Gamma(G,\Phi)$ has an edge between any two vertices, i.e., it is the complete graph K_n. In terms of optimization procedures and their move-sets, the complete graphs correspond to random search: all configurations are accessible in a single step.

Lemma 4. Let f be a non-constant landscape on the complete graph K_n with n vertices. Then f is elementary.

Elementary landscapes are thus only interesting when they "live" on non-trivial graphs with an interestingly rich spectrum of $-\Delta$.

Definition. The (Cartesian) product $\Gamma_1 \times \Gamma_2$ of two graphs has vertex set $V(\Gamma_1 \times \Gamma_2) = V(\Gamma_1) \times V(\Gamma_2)$. Two nodes (x_1,x_2) and (y_1,y_2) are connected if either (i) $x_1 = y_1$ and x_2, y_2 are adjacent in Γ_2, or (ii) $x_2 = y_2$ and x_1, y_1 are adjacent in Γ_1.

Proposition 3. Let $\Gamma = \Gamma_1 \times \Gamma_2$. Then

(i) Let $\lambda_k^{(1)}$ and $\lambda_j^{(2)}$ be eigenvalues of the Laplacians of the two graphs Γ_1 and Γ_2, respectively. Then λ is an eigenvalue of the Laplacian of $\Gamma_1 \times \Gamma_2$ if and only if it is of the form $\lambda_k^{(1)} + \lambda_j^{(2)}$.

(ii) Let $u_k^{(1)}$ and $u_j^{(2)}$ be eigenvectors of the Laplacian of Γ_1 and Γ_2. Then $u_k^{(1)} \times u_j^{(2)}$ is an eigenvector of $\Gamma_1 \times \Gamma_2$.

(iii) If $\Gamma_1 = \Gamma(G_1,\Phi_1)$ and $\Gamma_2 = \Gamma(G_2,\Phi_2)$ are Cayley graphs then

$$\Gamma(G_1,\Phi_1) \times \Gamma(G_2,\Phi) = \Gamma\Big(G_1 \times G_2, (\{\imath_1\} \times \Phi_2) \cup (\Phi_2 \times \{\imath_2\})\Big)$$

is again a Cayley graph.

Proof. For (i) and (ii) see [51], claim (iii) is easily verified from the definition of the Cartesian product. ∎

Our interest in the Cartesian product of graphs comes from the fact that important classes of graphs can be constructed as repeated Cartesian products of very simple units. As an example consider the *sequence spaces* or Hamming graphs $\mathcal{Q}_\alpha^n$. The vertices of these graphs are sequences of constant length n constructed from a fixed alphabet with α letters. Two sequences are adjacent if they differ in a single position along the sequence. Obviously $\mathcal{Q}_\alpha^1 \cong K_\alpha$ is the complete graph with α vertices. It is easy to check that $\mathcal{Q}_\alpha^n \cong \mathcal{Q}_\alpha^{n-1} \times K_\alpha$ for all $n \geq 2$.

4.2. Landscapes on the Boolean Hypercube

An orthonormal basis of eigenvectors of the Laplacian is easily constructed explicitly for Boolean hypercubes $\mathcal{Q}_2^n$. Without loosing generality we may use the alphabet $\{+1, -1\}$. A configuration is then a string σ of "spins" $\sigma_k \in \{+1, -1\}$. An alternative encoding uses a binary string x, with $x_i \in \{0, 1\}$. The following result is well known:

Proposition 4. Any landscape f on the Boolean Hypercube can be written as

$$f(\sigma) = J_0 + \sum_{p=1}^{n} \sum_{i_1 < i_2 < \ldots < i_p} J_{i_1 i_2 \ldots i_p} \sigma_{i_1} \sigma_{i_2} \ldots \sigma_{i_p} \,, \tag{13}$$

where the $J_{i_1 i_2 \ldots i_p}$ are constants.

It is not hard to check that the products

$$\varepsilon_q(\sigma) = \sigma_{i_1} \sigma_{i_2} \ldots \sigma_{i_p} \quad \text{where} \quad q_k = 1 \quad \text{if and only if} \quad k \in \{i_1, i_2, \ldots, i_p\}$$

are in fact eigenvector of the Laplacian of the Boolean hypercube because of the correspondence $\varepsilon_q(\sigma) = \chi_q(x)$ with $\sigma_k = 2x_k - 1$. Furthermore, one finds that the eigenvalue corresponding to ε_q depends only on the number p of non-zero entries in the multi-index q, see e.g. [34]. One finds

$$\lambda_p = 2np \quad \text{with multiplicity} \quad m(\lambda_p) = \binom{n}{p}. \tag{14}$$

Hamiltonians of the form

$$\mathcal{H}_p(\sigma) = \sum_{i_1<i_2<\ldots<i_p} J_{i_1 i_2 \ldots i_p} \sigma_{i_1} \sigma_{i_2} \ldots \sigma_{i_p} \tag{15}$$

play a prominent role in the theory of spin glasses; they are known as p-spin models. It was introduced by Derrida [31] in order to bridge the gap between the SK model [52], which is the special case $p = 2$, and the random energy model [31, 53, 32, 54]. In physics the coefficients are usually chosen i.i.d. from a Gaussian distribution. As a consequence of proposition 4 above we can represent any landscape on the Boolean hypercube as a superposition of p-spin models with suitable choices of the interaction coefficient $J_{i_1 i_2,\ldots,i_p}$. Let us now consider a few examples:

Weight Partition (`WP`). Given a string of n "spins" $x_i \in \{-1,+1\}^n$ and corresponding weights w_i, the cost function is given by

$$f(x) = \left(\sum_{i=1}^{n} w_i x_i \right)^2 . \tag{16}$$

The move set is given by flipping a single spin, hence the configuration space is again a hypercube. Grover [47] showed that `WP` is elementary with $\lambda = 4$ for any choice of the weights w_i.

Not-All-Equal-Satisfiability (`NAES`). Consider a vector of n binary variables. A literal is a variable or its complement. A clause is a set of three literals that does not contain both a variable and its complement. A clause is said to be satisfied if at least one literal is 0 and at least one literal is 1. An instance of `NAES` is given by a set of c clauses, and the cost function is the number of non-satisfied clauses. The move set is defined by flipping the value of a single variable, thus the configuration space is the Boolean Hypercube. Grover [47] showed that `WP` is elementary with $\lambda = 4$ for any choice of clauses.

Low Autocorrelated Binary Strings (LABSP). The LABSP [55, 56] consists of finding binary strings σ over the alphabet $\{-1,+1\}$ with low aperiodic off-peak autocorrelation $R_k(\sigma) = \sum_{i=1}^{N-k} \sigma_i\sigma_{i+k}$ for all lags k. These strings have technical applications such as the synchronization in digital communication systems and the modulation of radar pulses. The quality of a string σ is measured by the fitness function

$$f(\sigma) = \sum_{k=1}^{n-1} R_k(\sigma)^2. \tag{17}$$

In most of the literature on the LABSP the *merit factor* $F(\sigma) = n^2/(2f(\sigma))$ is used, see [56] for details.

Lemma 5. The landscape f of the LABSP can be written as

$$f(\sigma) = a_0 + \sum_{k=1}^{\lceil \frac{n}{2}\rceil -1} \sum_{i=1}^{n-2k} 2\varepsilon_{i,i+k}(\sigma) + \sum_{k=1}^{n-1}\sum_{i=1}^{n-1} \sum_{j\neq i-k,i,i+k} \varepsilon_{i,i+k,j,j+k}(\sigma). \tag{18}$$

Corollary 2. The "random walk" correlation function of the LABS is of the form

$$r(s) = [1-\mathcal{O}(1/n)]\left(1-\frac{8}{n}\right)^s + \mathcal{O}(1/n)\left(1-\frac{4}{n}\right)^s. \tag{19}$$

The landscape of the LABPS is thus not elementary, it consists of a superposition of two modes, namely $p=2$ and $p=4$. The smoother $p=2$ contribution becomes negligible for large n, so that f behaves for long strings almost like an elementary $p=4$ landscape. This fact explains why the LABPS has been found to be much harder for simulated annealing than, say, the SK spin glass [56]. This author [13] has computed the "random walk" autocorrelation function $r(s)$ numerically based on the merit factor F. The numerical estimate for the correlation length

$$\ell \approx 0.123 \cdot n - 0.983$$

is in excellent agreement with the asymptotic value $\ell = n/8 + \mathcal{O}(1)$ implied by the corollary.

4.3. Landscapes on Hamming Graphs

Boolean hypercubes are of course a special case of Hamming graphs. We have discussed them in a separate subsection because of their particular importance. Hamming graphs with larger alphabets ($\alpha > 2$) are on particular importance in biology: the sequences of nucleic acids, RNA or DNA, contain four different bases, and proteins use 20 different amino-acids. Just as for the hypercube, the Laplacian spectrum and an ONB of eigenvectors can be constructed explicitly [57, 58].

Graph Coloring Problem (GCP). An instance of a graph coloring problem consists of a graph $G(V,E)$ with n vertices and a collection of α different colors. A configuration x is an α-coloring of the vertex set V, i.e., an assignment of one color x_p to each vertex p of the graph. The cost function is the number edges $\{p,q\} \in E$ such that $x_p = x_q$:

$$f(x) = \sum_{\{p,q\}\in E} \delta_{x(p)x(q)}. \tag{20}$$

A move is the replacement of one color by another one at single vertex. The configuration spaces are thus the general Hamming graphs, i.e., sequence spaces over the alphabet of the α colors. Grover [47] has shown that each instance of GCP is elementary on $\mathcal{Q}_\alpha^n$ with $\lambda = 2\alpha$.

4.4. Permutations: Travelling Salesman and Graph Matching

The configurations of a family of optimization problems can be represented as permutations of finite number n of objects. Hence the symmetric group S_n takes the role of V. Natural choices of move sets are sets of generators of the S_n, and thus the configuration spaces are Cayley graphs of the symmetric group. The most convenient sets of generators are: the set $\mathcal{T}$ of all transpositions (i,j), the set $\mathcal{K}$ of all canonical transpositions $(i,i+1)$, and the set $\mathcal{J}$ of all reversals $[i,j]$, which are also called inversions or 2opt-moves, [59].

Travelling Salesman Problem (TSP). An instance of a TSP [9] is defined by a set of n cities and matrix W of costs for connecting them. A tour is a permutation π of cities, and its cost is

$$f(\pi) = \sum_{i=1}^{n} W_{\pi(i),\pi(i-1)} \tag{21}$$

where the indices are taken modulo n. Different versions of the TSP are defined by the properties of W (arbitrary, or symmetric, or with entries additionally obeying the triangle inequality, etc.; see [60]).

In the following it will be convenient to use

$$f^*(\pi) = f(\pi^*) = \sum_j w_{\pi(j)\pi(j+1)}.$$

The permutation π^* is the "reverse order" permutation of π, i.e., $f^*(\pi) = f(\pi^*)$ is the cost of the tour π when traveled in the opposite direction. Thus $f(\pi) = f^*(\pi)$ for all $\pi \in S_n$ is true if and only if the cost matrix W is symmetric. Recalling that any matrix W can be uniquely decomposed into its symmetric component $W^\sigma = (W + W^+)/2$ and its antisymmetric component $W^\alpha = (W - W^+)/2$ we introduce

$$f^\sigma(\pi) = \sum_j w^\sigma_{\pi(j)\pi(j-1)} = \frac{f(\pi) + f^*(\pi)}{2}, \quad f^\alpha(\pi) = \sum_j w^\alpha_{\pi(j)\pi(j-1)} = \frac{f(\pi) - f^*(\pi)}{2}.$$

Note that f^σ and f^α can be viewed as cost functions of TSPs with "distance matrices" W^σ and W^α, respectively.

Lemma 6. Both f^σ and f^α are elementary landscapes on the Cayley graphs of the symmetric group with the transpositions and the inversions as generators, respectively. In particular we have for transpositions

$$\Delta f^\sigma + 2(n-1)(f^\sigma - \bar{f}) = 0 \qquad \Delta f^\alpha + 2n f^\alpha = 0\,, \tag{22}$$

and for inversions (2opt moves [59]) we find

$$\Delta f^\sigma + n(f^\sigma - \bar{f}) = 0 \qquad \Delta f^\alpha + \frac{n(n+1)}{2} f^\alpha = 0\,. \tag{23}$$

Corollary 3. The landscape of a TSP with transpositions or inversions is elementary if and only if W is either symmetric or antisymmetric.

Asymmetric TSPs hence provide an example of fairly simple composite landscapes. They consist of two modes corresponding to the symmetric and the antisymmetric part of the distance matrix W, respectively. It is also interesting to note in this context that canonical transpositions $(i, i+1)$ do not lead to an elementary landscape. Numerical data [61, 11] have indicated that the "random walk" correlation functions $r(s)$ of both the symmetric and the antisymmetric components are exponential. Theorem 1 now provides a mathematical explanation for this observation.

The nearest neighbor correlations of the symmetric and antisymmetric components of a TSP with transpositions are $\varrho = 1 - 4/n$ and $\varrho = 1 - 4/(n-1)$, respectively, i.e., very similar. In fact, numerical estimates [11] are consistent with $\varrho \sim 1 - 4/n$ for large n in both cases. In the case of inversions we have a symmetric mode with nearest neighbor correlation $\varrho = 1 - 2/(n-1)$ and an antisymmetric contribution with a vanishingly small contribution $\varrho = -2/(n-1) \sim 0$.

It is interesting to correlate these values of ϱ with known facts about the performance of heuristic optimization algorithm, in particular with the simulated annealing. It has been observed by several authors that simulated annealing on symmetric TSP is much more effective when reversals instead of transpositions are used as move set, see, e.g., the books [62, 63]. Furthermore, Miller and Pekny [64] have observed that reversals are a remarkably inefficient move set for asymmetric `TSP`s. These observations are in accordance with the conjecture that landscape with smoother correlation functions have fewer local optima and are thus easier to optimize on [11]. In particular the difference between symmetric and asymmetric `TSP`s when reversals are used is easily explained in these terms: while for the symmetric `TSP` the landscape is as smooth as possible, it is extremely rugged for the antisymmetric case.

Graph Matching Problem (`GMP`). Given a graph G with n vertices and a symmetric matrix W of edge weights, the task is to partition the graph into $n/2$ pairs of vertices such that the sum of the edge weights corresponding to these pairs is optimal. A convenient encoding of the problem is the following. Let $\pi \in S_n$ be a permutation of the vertices. We assume that the vertices are arranged such that $[\pi(2k-1), \pi(2k)]$ form a pair. The cost function is then

$$f(\pi) = \sum_{k=1}^{n/2} W_{\pi(2k-1),\pi(2k)} \,. \tag{24}$$

Again, the configuration space is the symmetric group, and hence the set of all transpositions is a reasonable move set.

Lemma 7. The landscape of the graph matching problem `GMP` with transposition metric is elementary with $\lambda = 2(n-1)$.

Note that this result is false if W is not symmetric.

4.5. Landscapes on Johnson Graphs

Table 1. Summary of Elementary Landscapes.

Problem	Γ	D	λ	ϱ	ℓ
`NAES`	$\mathcal{Q}_2^n$	n	4	$1-\frac{4}{n}$	$\frac{1}{4}n-\frac{1}{2}-\frac{1}{3}\frac{1}{n}+\mathcal{O}(\frac{1}{n^2})$
`WPP`	$\mathcal{Q}_2^n$	n	4	$1-\frac{4}{n}$	$\frac{1}{4}n-\frac{1}{2}-\frac{1}{3}\frac{1}{n}+\mathcal{O}(\frac{1}{n^2})$
p-spin	$\mathcal{Q}_2^n$	n	$2p$	$1-\frac{2p}{n}$	$\frac{1}{2p}n-\frac{1}{2}-\frac{p}{6}\frac{1}{n}+\mathcal{O}(\frac{1}{n^2})$
`GCP`	$\mathcal{Q}_\alpha^n$	$(\alpha-1)n$	2α	$1-\frac{2\alpha}{(\alpha-1)n}$	$\frac{\alpha-1}{2\alpha}n-\frac{1}{2}+\frac{\alpha}{6(\alpha-1)}\frac{1}{n}+\mathcal{O}(\frac{1}{n^2})$
symmetric	$\Gamma(S_n,\mathcal{T})$	$n(n-1)/2$	$2(n-1)$	$1-\frac{4}{n}$	$\frac{1}{4}n-\frac{1}{2}-\frac{1}{3}\frac{1}{n}+\mathcal{O}(\frac{1}{n^2})$
`TSP`	$\Gamma(S_n,\mathcal{J})$	$n(n-1)/2$	n	$1-\frac{2}{n-1}$	$\frac{1}{2}n-1-\frac{1}{6}\frac{1}{n}+\mathcal{O}(\frac{1}{n^2})$
anti-symmetric	$\Gamma(S_n,\mathcal{T})$	$n(n-1)/2$	$2n$	$1-\frac{4}{n-1}$	$\frac{1}{4}n-\frac{3}{4}-\frac{1}{3}\frac{1}{n}+\mathcal{O}(\frac{1}{n^2})$
`TSP`	$\Gamma(S_n,\mathcal{J})$	$n(n-1)/2$	$n(n+1)/2$	$-\frac{2}{n-1}$	$\frac{1}{\ln n}-\ln 2\frac{1}{\ln^2 n}+\mathcal{O}(\ln^{-3} n)$
`GMP`	$\Gamma(S_n,\mathcal{T})$	$n(n-1)/2$	$2(n-1)$	$1-\frac{4}{n}$	$\frac{1}{4}n-\frac{1}{2}-\frac{1}{3}\frac{1}{n}+\mathcal{O}(\frac{1}{n^2})$
`GBP`	$J(n,n/2)$	$n^2/4$	$2(n-1)$	$1-\frac{8}{n}+\frac{8}{n^2}$	$\frac{1}{8}n-\frac{3}{8}-\frac{13}{24}\frac{1}{n}+\mathcal{O}(\frac{1}{n^2})$

Another class of configuration spaces arises if the configurations can be regarded as subsets of some finite set.

Definition. Let X be a finite set, $n = |X|$, and let $\mathcal{S}_k$ be the collection of k-element subsets of X. The graph $J(n,k)$ has vertex set $\mathcal{S}_k$ and two vertices are adjacent if the corresponding subsets of X have $k-1$ vertices in common. $J(n,k)$ is called *Johnson graph.*

Only one example of this class has received extensive attention so far.

Graph Bipartitioning Problem (GBP). G is a graph with an even number n of vertices and H is a symmetric matrix of edge weights. A configuration is a partition of the vertex set into two subsets A and $X \setminus A$ of equal size. Two bi-partitions $[A, X \setminus A]$ and $[B, X \setminus B]$ are neighbors if B is obtained from A by exchanging a vertex from A by a vertex from $X \setminus A$. Thus the configuration space is the Johnson graph $J(n, n/2)$. The cost function is

$$f([A, X \setminus A]) = \sum_{i \in A} \sum_{j \in X \setminus A} H_{ij} \, , \tag{25}$$

i.e., f is the total weight of all edges connecting A and $X \setminus A$. As a close relative of the Sherrington-Kirkpatrick model the GBP has received considerable attention [10]. Grover [47] found that each instance of the GBP is an elementary landscape with $\lambda = 2(n-1)$, see also table 2. Stadler and Happel [23] have shown by explicit calculations that the "random walk" correlation function of the GBP on $J(n, n/2)$ is

$$r(s) = \left(1 - \frac{8}{n} + \frac{8}{n^2}\right)^s \, , \tag{26}$$

in accordance with Theorem 1.

Table 1 summarizes the known elementary landscapes.

4.6. Nodal Domains of Schrödinger Operators

In this contribution we are only interested in eigenfunctions φ of the Laplacian operator $-\Delta$. The following interesting result can, however, be proved for a much larger class of operators acting on landscapes. Let a be a weight function on the edges of Γ, conveniently defined as $a : V\times V \to \mathbb{R}$ such that $a(x,y) = a(y,x) > 0$ if $\{x,y\} \in E$ and $a(x,y) = 0$ otherwise. Furthermore let $v : V \to \mathbb{R}$ be an arbitrary "potential". A linear operator $\mathbf{H}$, defined by the action

$$\mathbf{H}f(x) \stackrel{\text{def}}{=} \sum_{x\sim y} a(x,y)\,[f(x) - f(y)] + v(x)f(x)\,, \tag{27}$$

is called a *Schrödinger operator*. Setting $a(x,y) = 1$ if $\{x,y\} \in E$, i.e., $a(x,y) = \mathbf{A}_{xy}$, and choosing $v(x) = 0$ shows that the graph Laplacian $-\Delta$ is in fact a Schrödinger operator. Discrete Schrödinger operators without potential can be interpreted as Laplacians of edge-weighted graphs [65].

The Perron-Frobenius theorem implies that the smallest eigenvalue λ_1 of $\mathcal{H}$ is non-degenerate and the corresponding eigenfunction f_1 is positive everywhere if Γ is connected. Let

$$\lambda_1 < \lambda_2 \le \lambda_3 \le \ldots \le \lambda_{k-1} \le \lambda_k \le \lambda_{k+1} \le \ldots \le \lambda_{|V|}$$

be the list of eigenvalues of $\mathbf{H}$ arranged in non-decreasing order and repeated according to multiplicity[3]. Let f_k be *any* eigenfunction associated with the eigenvalue λ_k. Without loosing generality we may assume that $\{f_i\}$ is a complete orthonormal set of eigenfunctions satisfying $\mathbf{H}f_i = \lambda_i f_i$. Since $\mathbf{H}$ is a real operator, we take all eigenfunctions to be real.

A configuration $x \in V$ is a *local minimum* if $f(x) \le f(y)$ for all neighbors $y \in \partial\{x\}$ of x. Correspondingly x is a *local maximum* if $f(x) \ge f(y)$ for all $y \in \partial\{x\}$. Local optima are the very feature of a landscape that makes it rugged.

[3]Note that in this subsection we count the eigenvalues starting at 1 instead of 0 as in the rest of this contribution

Local optima are traps of optimization heuristics and evolutionary adaptation. An understanding of the distribution of local optima is thus of utmost importance for the understanding of a landscape. Little can be said about the geometric arrangement of local optima for an arbitrary landscape f. The situation is slightly better for elementary landscapes.

Proposition 5. Let f be an eigenfunction of $\mathbf{H}$ with eigenvalue $\lambda > 0$. If x_0 is a local maximum of f then $v(x_0) \leq \lambda$ implies $f(x_0) \geq 0$ and $v(x_0) \geq \lambda$ implies $f(x_0) \leq 0$. If x_0 is a local minimum of f then $v(x_0) \leq \lambda$ implies $f(x_0) \leq 0$ and $v(x_0) \geq \lambda$ implies $f(x_0) \geq 0$. In particular, if $v \equiv 0$ then all local maxima are non-negative and all local minimal are non-positive.

Proof. If x_0 is local maximum then $f(x_0) - f(y) \geq 0$, and hence $\mathbf{H}f(x_0) \geq v(x_0)f(x_0)$. Using that f is an eigenfunction yields $\lambda f(x_0) \geq v(x_0)f(x_0)$. Using that $\lambda \geq 0$ yields the assertion for a local maximum. The argument for local minima is analogous. If $v \equiv 0$, i.e., if $\mathbf{H}$ is a Laplacian of an edge-weighted graph, the case $\lambda \leq v(x_0)$ contradicts $\lambda > 0$. The special case $\mathbf{H} = -\Delta$ is theorem 6 in [47]. ∎

Some more global information can be obtained on the distribution of positive and negative valued of an eigenfunction f of a Schrödinger operator. This leads us to the notion of nodal domains. In a continuous setting one defines the nodal set of a continuous function f as the preimage $f^{-1}(0)$. The nodal domains are the connected components of the complement of $f^{-1}(0)$. In the discrete case this definition does not make sense since a function f can change sign without having zeroes. Instead we use the following

Definition. X is a nodal domain of a function $f : V \to \mathbb{R}$ if it is a maximal subset of V subject to the two conditions

(i) X is connected as an induced subgraph of Γ;

(ii) if $x, y \in X$ then $f(x)f(y) \geq 0$.

The following properties of nodal domains of an eigenfunction f of a Schrödinger operator can be easily verified.

(a) Every point $x \in V$ lies in some nodal domain $X \subset V$.

(b) If V is a nodal domain then it contains at least one point $x \in V$ with $f(x) \neq 0$ and f has the same sign on all non-zero points in X. Thus each nodal-domain can be called either "positive" or "negative".

(c) If two nodal domains X and X' have non-empty intersection the $f|_{X \cap X'} = 0$ and X, X' have opposite sign or $X = X'$.

The following result generalizes Courant's nodal domain theorem for Riemannian manifolds, see [66], to arbitrary connected graphs.

Proposition 6. The eigenfunction f_k has at most k nodal domains.

Proof. See [67]. ∎

Remark. The second-largest eigenvalue of Graph, often called the *algebraic connectivity*, and its eigenvectors, which are sometimes referred to as *characteristic valuations*, have received some attention in graph theory [68, 69, 70]. The case $k = 2$ of proposition 5 was proved already in 1975 by Fiedler [71]. Stuart Kauffman [72] calls this type of landscape "Fujijama", because they have only a single mountain massive (positive nodal domain).

5. Autocorrelation Functions and Coherent Configurations

5.1. Autocorrelation Functions on Partitions

The use of "random walk" correlation functions for characterizing landscapes has two disadvantages: (i) the information on the landscape is "blurred" by the transition matrix $\mathbf{T}$ of the walk, and (ii) samples along a random walk converge slowly compared to sampling independent pairs of configurations. For these reasons most of the results on the correlation structure of RNA landscapes, where the evaluation of f is extremely costly in terms of computer resources, have been obtained in terms of the correlation functions $\rho(d)$, which is defined as the average correlation of all pairs (x, y) of vertices with distance $d(x, y) = d$ [15, 16, 24]. In this section we link this approach to the theory of the "random walk" correlation functions developed above.

It is convenient to formulate our discussion in terms of partitions of the set of (ordered) pairs of vertices $V \times V$. Recall that there is a canonical metric on the vertex set V of any connected graph Γ. The distance $d : V \times V \to \mathbb{N}_0^+$ on V is defined as the minimum number of edges separating two vertices. If Γ is finite, then there is a maximum distance, which is called the *diameter* d^* of Γ. The properties of the distance on V are discussed in detail in the book [73]. The metric $d(\,.\,,\,.\,)$ induces in a natural way the *distance partition* $\mathfrak{D}$ of $V \times V$ which has the classes

$$\mathcal{D}_d \stackrel{\text{def}}{=} \left\{ (x, y) \in V \middle| d(x, y) = d \right\}. \tag{28}$$

The distance partition seems to be the most natural and useful partition of $V \times V$. The subsequent discussion will show, however, that this is true only for sufficiently "symmetric" graphs Γ. For the moment we consider arbitrary partitions of $V \times V$, postponing the problem of choosing the partitions until the following subsections.

Definition. Let $\mathfrak{R}$ be a partition of $V \times V$ and suppose $f : V \to \mathbb{R}$ is non-constant. Then the *correlation function* $\rho : \mathfrak{R} \to [-1, 1]$ of f with respect to $\mathfrak{R}$ is defined by

$$\rho(\mathcal{X}) \stackrel{\text{def}}{=} \frac{1}{|\mathcal{X}| \cdot \sigma^2} \sum_{(x,y)\in\mathcal{X}} \left[(f(x) - \overline{f})(f(y) - \overline{f}) \right] \tag{29}$$

where $\mathcal{X}$ is a class of the partition $\mathfrak{R}$.

Let us note a few general properties of correlation functions on partitions. A matrix $\mathbf{X}$ with entries

$$\mathbf{X}_{xy} \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } (x, y) \in \mathcal{X} \\ 0 & \text{otherwise} \end{cases} \tag{30}$$

is associated with each class $\mathcal{X} \in \mathfrak{R}$. In order to simplify the formalism in the following we will always assume that $\overline{f} = 0$. Since both $r(s)$ and $\rho(\mathcal{X})$ are invariant under the transformation $f(x) \to f(x) - \overline{f}$ we can do this without loosing generality. Thus we have

$$\rho(\mathcal{X}) = \frac{|V|}{|\mathcal{X}|} \frac{\langle f, \mathbf{X}f \rangle}{\langle f, f \rangle}\,. \tag{31}$$

Lemma 8. Let $\mathfrak{R}$ be a partition of $V \times V$ and let f be a non-constant landscape on V. Then $\sum_{\mathcal{X}\in\mathfrak{R}} \rho(\mathcal{X}) \cdot |\mathcal{X}| = 0$.

Lemmarl-0 means that there is no "average correlation" in a sample of random points, because $|\mathcal{X}|/|V|^2$ is the probability that two randomly picked points x and y form a pair $(x, y) \in \mathcal{X}$.

The *diagonal* of $V \times V$ is $\mathcal{I} \stackrel{\text{def}}{=} \{(x,x) \,|\, x \in V\}$; the associated matrix is the identity matrix $\mathbf{I}$. The transpose of a class $\mathcal{X} \in \mathfrak{R}$ is

$$\mathcal{X}^+ \stackrel{\text{def}}{=} \{(y,x) \,|\, (x,y) \in \mathcal{X}\} \ .$$

If $\mathbf{X}$ is the matrix associated with the class $\mathcal{X}$, then its transpose $\mathbf{X}^+$ is associated with the class $\mathcal{X}^+$. The definition of ρ immediately implies

$$\rho(\mathcal{X}) = \rho(\mathcal{X}^+) \qquad \text{and} \qquad \rho(\mathcal{I}) = 1\,. \tag{32}$$

The following two conditions seem to be quite natural for our purposes:

Definition. Let $\mathfrak{R}$ be a partition of $V \times V$. Following Higman [74] call $\mathfrak{R}$ a *pre-coherent configuration* if the implication $\mathcal{X} \cap \mathcal{I} \neq \emptyset \implies \mathcal{X} \subseteq \mathcal{I}$ holds and if for any $\mathcal{X} \in \mathfrak{R}$ we have also $\mathcal{X}^+ \in \mathfrak{R}$.

The symmetry of ρ under transposition might suggest to require $\mathcal{X} = \mathcal{X}^+$ and $\mathcal{I} \in \mathfrak{R}$. For technical reasons (which will become clear later on) it is wise to abstain from these more stringent conditions. Since we are given not only the set V but also the neighborhood structure of Γ it seems natural to require that it is in some way respected by the partition $\mathfrak{R}$. The edge set E of Γ translates into to the set

$$\mathcal{A} \stackrel{\text{def}}{=} \{(x,y) \,|\, \{x,y\} \in E\} = \mathcal{D}_1. \tag{33}$$

of ordered pairs of vertices. The matrix associated with $\mathcal{A}$ is of course the adjacency matrix $\mathbf{A}$ of Γ.

Definition. A partition $\mathfrak{R}$ of $V \times V$ *compatible* with Γ if $\mathcal{X} \cap \mathcal{A} \neq \emptyset$ implies $\mathcal{X} \subseteq \mathcal{A}$ for all $\mathcal{X} \in \mathfrak{R}$.

Unfortunately, pre-coherent configurations compatible with Γ are still too general to allow for an interesting theory. In the following section we will therefore introduce one more condition on $\mathfrak{R}$.

5.2. Coherent Configurations

Definition. A pre-coherent configuration $\mathfrak{R}$ is *coherent* if for all classes $\mathcal{X},\mathcal{Y},\mathcal{Z}\in\mathfrak{R}$ the following statement is true:
The numbers $\big|\{z\in V\,\big|\,(x,z)\in\mathcal{X}\text{ and }(z,y)\in\mathcal{Y}\}\big| \stackrel{\text{def}}{=} p^{\mathcal{Z}}_{\mathcal{X}\mathcal{Y}}$ are the same for all pairs $(x,y)\in\mathcal{Z}$.
The numbers $p^{\mathcal{Z}}_{\mathcal{X}\mathcal{Y}}$ are called the *intersection numbers* of $\mathfrak{R}$.

Coherent configurations have been studied in detail by Higman [74, 75, 76]. The highlights of his theory will be outlined later on in this subsection.

Let $\mathfrak{R}=\{\mathcal{R}_1,\mathcal{R}_2,\dots\mathcal{R}_r\}$ and $\mathfrak{S}=\{\mathcal{S}_1,\mathcal{S}_2,\dots\mathcal{S}_s\}$ be two partitions of a set W. We say that $\mathfrak{R}$ is a *refinement* of $\mathfrak{S}$ if for all $k\in[1,r]$ there is a $j\in[1,s]$ (depending on k) such that $\mathcal{R}_k\subseteq\mathcal{S}_j$. The number of classes in a partition is called its rank. We write $\mathfrak{S}\preceq\mathfrak{R}$ because the ranks fulfill $|\mathfrak{S}|\le|\mathfrak{R}|$.

The partition $\mathfrak{M}$ of $V\times V$ with classes $\{(x,y)\}$ for all $x,y\in V$ is called the maximum configuration, and the partition $\mathfrak{N}$ consisting of the two classes $\mathcal{I}$ and $\mathcal{K}=\{(x,y)|x\neq y\}$ is called the minimum configuration of $V\times V$. It is easy to check that both $\mathfrak{M}$ and $\mathfrak{N}$ are coherent. The following result provides a useful condition for compatibility with Γ:

Lemma 9. A coherent configuration $\mathfrak{R}$ on V is compatible with Γ if and only if $\mathfrak{R}$ is a refinement of the distance partition $\mathfrak{D}$ on $V\times V$, i.e., $\mathfrak{D}\preceq\mathfrak{R}$.

Let Γ be an arbitrary graph with vertex set V, then the maximum configuration $\mathfrak{M}$ on V is consistent with Γ. Of course $\mathfrak{M}$ is not interesting in itself because it does not contain any information about Γ. It guarantees, however, that the theory developed in the remainder of this contribution is well defined for all graphs. Clearly, our considerations will be of interest only if Γ admits compatible coherent configurations that are much coarse than $\mathfrak{M}$. It will be shown in the following section that the most interesting landscapes in fact "live" on configuration spaces that admit very coarse coherent configurations. The only graph with which the minimum configuration $\mathfrak{N}$ is compatible, however, is the complete graph $K_{|V|}$ with vertex set V.

The importance of coherent configurations comes from the algebraic properties of the matrices associated with its classes.

Proposition 7. The partition $\mathfrak{R}$ of $V \times V$ is a coherent configuration if and only if the matrices associated with the classes of $\mathfrak{R}$ fulfill:

(i) $\sum_{\mathcal{X} \in \mathfrak{R}} \mathbf{X} = \mathbf{J}$.

(ii) $\mathbf{I}$ is the sum of some elements of $\mathfrak{R}$.

(iii) $\mathbf{X} \in \mathfrak{R}$ implies $\mathbf{X}^+ \in \mathfrak{R}$

(iv) $\mathbf{XY} = \sum_{\mathcal{Z} \in \mathfrak{R}} p^{\mathcal{Z}}_{\mathcal{X}\mathcal{Y}} \mathbf{Z}$

Proof. See [74, 76]. ∎

The matrices associated with the classes of $\mathfrak{R}$ form therefore (the standard basis of) an algebra $\langle\mathfrak{R}\rangle$ that is closed (i) under the component-wise (i.e. Schur or Hadamard) product, (ii) closed under ordinary matrix multiplication and (iii) closed under transposition. Conversely, any matrix algebra fulfilling (i), (ii), and (iii) has a standard basis of the above form, i.e., there is a one-to-one correspondence between coherent configurations and matrix algebras of this type, for which Higman [76] has introduced the term *coherent algebra*. The condition for compatibility with Γ translates to "$\langle\mathfrak{R}\rangle$ is compatible with Γ if and only if the adjacency matrix $\mathbf{A}$ is contained in $\langle\mathfrak{R}\rangle$".

Proposition 8. The matrices defined by $\hat{\mathbf{Y}}_{\mathcal{X}\mathcal{Z}} = p^{\mathcal{Z}}_{\mathcal{X}\mathcal{Y}}$ for all $\mathcal{X}, \mathcal{Y}, \mathcal{Z} \in \mathfrak{R}$ form a matrix algebra $\langle\hat{\mathfrak{R}}\rangle$, the so-called *intersection algebra*. $\langle\hat{\mathfrak{R}}\rangle$ is isomorphic to $\langle\mathfrak{R}\rangle$.

Proof. See [74, 76]. ∎

It will sometimes be necessary to consider additional conditions on $\mathfrak{R}$. The most important ones have been considered already in Higman's papers:

Definition. A coherent configuration $\mathfrak{R}$ is

- *homogeneous* if $\mathcal{I} \in \mathfrak{R}$;
- *commutative* if $p^{\mathcal{Z}}_{\mathcal{X}\mathcal{Y}} = p^{\mathcal{Z}}_{\mathcal{Y}\mathcal{X}}$ for all $\mathcal{X}, \mathcal{Y}, \mathcal{Z} \in \mathfrak{R}$, i.e., if $\mathbf{XY} = \mathbf{YX}$ for all $\mathbf{X}, \mathbf{Y} \in \langle\hat{\mathfrak{R}}\rangle$;
- *symmetric* if $\mathcal{X} = \mathcal{X}^{+}$ for all $\mathcal{X} \in \mathfrak{R}$.

Higman [74] showed that symmetry implies commutativitity which in turn implies homogeneity. Conversely, a homogeneous coherent configuration with rank $|\mathfrak{R}| \leq 5$ is already commutative [74, (4.1)]. Commutative coherent configurations have important applications in coding theory [77, 78]. They have been studied extensively under the name *association schemes* by Delsarte [79]. Symmetric coherent configurations are oftentimes termed *symmetric association schemes* or simply association schemes; the matrix algebra associated with them is known as *Bose-Mesner algebra* [80].

The adjacency matrix $\mathbf{A}$ of a graph Γ generates an algebra $\langle\mathfrak{A}\rangle$, the so-called *adjacency algebra* of Γ, see e.g. [43]. Now let $\mathfrak{R}$ be a coherent configuration compatible with Γ. Since $\mathbf{A} \in \mathfrak{R}$ it is clear that the adjacency algebra is a subalgebra of the coherent algebra $\langle\mathfrak{R}\rangle$. It has been shown [74] that $\langle\mathfrak{A}\rangle = \langle\mathfrak{R}\rangle$ if and only if $\mathfrak{R}$ is commutative. This result can be used to explicitly construct $\mathfrak{R}$ for a given graph, see [81].

Not all graphs admit interestingly small coherent configurations. On the other hand there are interesting classes of graphs that give rise to very coarse coherent configurations. A prominent example are *distance regular* graphs, which have attracted a lot of interest in discrete mathematics, as the recent monograph [82] shows. As an immediate consequence of their definition we have the following

Remark. Γ is distance regular if and only if its distance partition $\mathfrak{D}$ is a coherent configuration. Examples include the Hamming graphs and the Johnson graphs but not the Cayley graphs of the symmetric group discusses in the previous section.

In the following subsection we will briefly review a very general construction for coherent configurations. All this material is known in the literature, it will serve, however, as an additional motivation for the application of coherent configurations to the study of landscapes.

5.3. Graph Automorphisms

An automorphism of the graph $\Gamma = (V, E)$ is a one-to-one map $\alpha : V \to V$ such that $\alpha(x)$ and $\alpha(y)$ are adjacent if and only if x and y are adjacent. The set of all automorphisms of a graph forms a group under composition, the so-called automorphism group $\mathbf{Aut}[\Gamma]$. It is a permutation group acting on the point set V.

Now consider a permutation group G acting on a finite set V. On $V \times V$ is acts component-wise: $\alpha(x,y) \stackrel{\text{def}}{=} \big(\alpha(x), \alpha(y)\big)$. The orbits of G on $V \times V$ are called *orbitals*. They form a partition $\mathfrak{G}$ on $V \times V$.

Proposition 9. The partition $\mathfrak{G}$ of $V \times V$ induces by the orbitals of a group G acting on V is a coherent configuration. The coherent algebra $\langle \mathfrak{G} \rangle$ coincides with the centralizer algebra of the permutation representation of the group G.

Proof. See [83, 84, 74]. ∎

Thus any graph Γ with a non-trivial automorphism group $\mathbf{Aut}[\Gamma]$ admits a coherent configuration $\mathfrak{G} \prec \mathfrak{M}$ that is strictly coarser than the maximum configuration $\mathfrak{M}$. Important properties of the permutation groups translate into the properties of coherent configurations discussed at the end of the previous section:

Proposition 10. Let G be a permutation group acting on V and let $\mathfrak{G}$ be the coherent configuration induced on $V \times V$. Then:

$$\begin{array}{lllcccc} G & \text{is} & \text{transitive} & \Longleftrightarrow & \mathfrak{G} & \text{is} & \text{homogeneous} \\ G & \text{is} & \text{multiplicity-free} & \Longleftrightarrow & \mathfrak{G} & \text{is} & \text{commutative} \\ G & \text{is} & \text{generously transitive} & \Longleftrightarrow & \mathfrak{G} & \text{is} & \text{symmetric} \end{array}$$

Proof. See [84]. ∎

The first condition is of particular importance: All Cayley graphs have transitive groups of automorphisms [85], and thus the corresponding coherent configurations $\mathfrak{G}$ are homogeneous. The Hamming graphs and the Johnson graphs, for example, are distance transitive, i.e., for any two pairs of vertices (u, v) and (x, y) fulfilling $d(x, y) = d(u, v)$ there is an automorphism α such that $\alpha(x, y) = (u, v)$, see, e.g., [82]. This implies in particular that $\mathbf{Aut}[\Gamma]$ is generously transitive, i.e., that for any two vertices $u, v \in V$ there is an automorphism fulfilling $\alpha(u, v) = (v, u)$.

5.4. Equitable Partitions

Definition. Let ϖ be a partition of the vertex set V of Γ. ϖ is called *equitable* if the number of neighbors which a vertex $y \in Y$ has in class X is independent of the choice of the vertex in Y. In other words, ϖ is equitable if for all $X, Y \in \varpi$ holds

$$\hat{\mathbf{A}}_{XY} \stackrel{\text{def}}{=} |\partial\{y\} \cap X| = \sum_{x \in X} A_{xy} \qquad \text{for all } y \in Y. \tag{34}$$

We call $\hat{\mathbf{A}}$ the *collapsed adjacency matrix* of Γ with respect to ϖ. If an equitable partition ϖ contains a class $\varpi_0 = \{u\}$ consisting of a single vertex $u \in V$ we will say that it is *anchored at (the reference vertex) u.*

With each partition ϖ of V into $M+1$ classes there is an associated $(M+1) \times |V|$ matrix, which we will also denote by ϖ. Its entries are

$$\varpi_{Xx} \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } x \in X \\ 0 & \text{otherwise} \end{cases} \tag{35}$$

We remark that this definition is the transpose of the convention in Godsil's book [77], while it conforms the notation in Bollobás' book [86]. Equitable partitions have been introduced by Schwenk [87]; more recently they have been used by Powers and coworkers under the name *colorations*, see, e.g., [88, 89], see also [51, Chap.4]. The following lemma collects their most useful properties.

Proposition 11. Let ϖ an equitable partition of the vertex set V of a graph Γ. Then the following statements hold.

(i) $\varpi\mathbf{A} = \hat{\mathbf{A}}\varpi$.

(ii) If φ is an eigenvector of A with eigenvalue μ, then $v = \varpi\varphi$ is a *right* eigenvector of $\hat{\mathbf{A}}$ with the same eigenvalue μ provided $v \neq 0$. For later reference we note the explicit formula

$$v(X) = \sum_{x\in X} \varphi(x)\,. \tag{36}$$

(iii) If v is a right eigenvector of $\hat{\mathbf{A}}$, then $u(X) = (1/|X|)v(X)$ defines a left eigenvector of $\hat{\mathbf{A}}$.

(iv) If u is a left eigenvector of $\hat{\mathbf{A}}$, then $u\varpi$ is an eigenvector of $\mathbf{A}$ which is constant on all classes of ϖ.

(v) $|X|[\hat{\mathbf{A}}^s]_{XY} = |Y|[\hat{\mathbf{A}}^s]_{YX}$ for all s.

(vi) The characteristic polynomial of $\hat{\mathbf{A}}$ divides the characteristic polynomial of $\mathbf{A}$.

(vii) If Γ is connected then $\mathbf{A}$ and $\hat{\mathbf{A}}$ have the same spectral radius.

(vi') If ϖ is anchored at some reference vertex $x_0 \in V$ then $\mathbf{A}$ and $\hat{\mathbf{A}}$ have the same minimal polynomial.

(vii') If ϖ is anchored at some reference vertex $x_0 \in V$ then μ is an eigenvalue of $\hat{\mathbf{A}}$ if and only if μ is an eigenvalue of $\mathbf{A}$.

(viii) $p(\hat{\mathbf{A}}) = \hat{p}(\mathbf{A})$ for any polynomial p.

(ix) If ϖ is anchored at $x_0 \in V$ then $[\mathbf{A}^s]_{yx_0} = [\hat{\mathbf{A}}^s]_{y'x_0}$ whenever y and y' belong to the same class of ϖ.

Proof. See [77, chap.5] for the proofs of (i) through (vii). Properties (vi') and (vii') are [86, Thm.8.6], (viii) is [90, Lem.1], and (ix) follows from (viii) and $[\mathbf{A}^s]_{yx_0} =$

$[\mathbf{A}^s]_{x_0 y} = [\hat{\mathbf{A}}^s]_{x_0 Y}$ for all $y \in Y$ which is true because $\{x_0\}$ is a class of ϖ on its own. ∎

Definition. Let $\mathfrak{R}$ be a partition of $V \times V$, and let $x_0 \in V$. Then the partition $\mathfrak{R}_{x_0}$ of V *anchored at (the reference vertex)* $x_0 \in V$ has the classes

$$\mathcal{X}_{x_0} = \{x \in V \,|\, (x, x_0) \in \mathcal{X}\} \,, \tag{37}$$

where $\mathcal{X} \in \mathfrak{R}$.

It is trivial to check that $\mathfrak{R}_{x_0}$ is in fact a partition of V. Note that $\mathcal{X}_{x_0}$ as defined above can be empty. The interest in this constructions originates from the following result:

Lemma 10. Let $\mathfrak{R}$ be a coherent configuration on V and let $x_0 \in V$. Then $\mathfrak{R}_{x_0}$ is an equitable partition of V anchored at the reference vertex x_0. The entries of the collapsed adjacency matrix fulfill

$$\hat{\mathbf{A}}_{\mathcal{X}_{x_0}\mathcal{Y}_{x_0}} = \sum_{\substack{\mathcal{Z} \in \mathfrak{A} \\ \mathcal{Z} \cap \mathcal{A} \neq \emptyset}} p^{\mathcal{Z}}_{\mathcal{X}\mathcal{Y}} \,. \tag{38}$$

By equ.(38) the entries in $\hat{\mathbf{A}}$ do not depend explicitly on the reference vertex x_0. If $\mathfrak{R}$ is homogeneous, then none of the classes $\mathcal{X}_{x_0}$ in the "projection" is empty, and we can assume that the rows and columns of $\hat{\mathbf{A}}$ are indexed by the classes of the coherent configuration $\mathfrak{R}$.

Note that the collapsed adjacency matrix $\hat{\mathbf{A}}$ obtained from a coherent configuration $\mathfrak{R}$ as described above is contained in the intersection algebra $\langle\hat{\mathfrak{R}}\rangle$. In fact, a basis of $\langle\hat{\mathfrak{R}}\rangle$ is obtained by the same procedure:

Proposition 12. Let $\mathfrak{R}$ be a coherent configuration and $x_0 \in V$. Then

$$p^{\mathcal{Z}}_{\mathcal{X}\mathcal{Y}} = \hat{\mathbf{Y}}_{\mathcal{X}\mathcal{Z}} = \sum_{x \in \mathcal{X}_{x_0}} \mathbf{Y}_{xz} \quad \text{for any } z \in \mathcal{Z} \,. \tag{39}$$

Proof. See [74]. ∎

Proposition 13. Let $\mathfrak{R}$ be a coherent configuration compatible with Γ. Then $\langle\mathfrak{R}\rangle$ is commutative if and only if all eigenvalues of $\hat{\mathbf{A}}$ are simple.

Proof. See [74]. ∎

The following proposition is the main result of [91]. In this paper ϖ has been assumed to arise from the orbits of a transitive automorphism group of Γ. However, only the properties of equitable partitions anchored at a reference vertex $x_0 \in V$ are actually used for its proof.

Proposition 14. Let Γ be a D-regular connected graph with adjacency matrix $\mathbf{A}$ and let ϖ be an equitable partition of V anchored at $x_0 \in V$. Let $g : \varpi \to \mathbb{R}$ be a real valued function of the classes of ϖ and define $g^* : \mathbb{N} \to \mathbb{R}$ by

$$g^*(s) \stackrel{\text{def}}{=} \sum_{Y\in\varpi} \left(\sum_{y\in Y} [\mathbf{T}^s]_{yx_0} \right) g(Y) . \tag{40}$$

Then g^* is an exponential function if and only if g is a *left* eigenvector of the collapsed adjacency matrix $\hat{\mathbf{A}}$ of Γ. In this case $g^*(s) = g(\{x_0\}) \times (\mu/D)^s$.

Proof. We sketch only the main ideas of the proof here. The details can be found in [91].
(i) The first step is to construct bases of left eigenvectors $\{u_i\}$ and right eigenvectors $\{v_i\}$ of $\hat{\mathbf{A}}$ that fulfill $\langle u_i, v_j\rangle = c_i\delta_{ij}$ with constant $c_i > 0$. This is possible since $\hat{\mathbf{A}}$ is diagonalizable. One can show furthermore that one may choose these bases such that $u(\{x_0\}) = v(\{x_0\}) = \varphi(x_0) = 1/\sqrt{|V|}$.
(ii) The next step is to express $[\mathbf{T}^s]_{yx_0}$ in terms of the collapsed adjacency matrix. One finds explicitly

$$\sum_{y\in Y} [\mathbf{T}^s]_{yx_0} = \frac{1}{\sqrt{|V|}} \sum_k \frac{1}{c_k} \left(\frac{\mu_k}{D}\right)^s v_k(Y).$$

(iii) The final step consists in expanding g w.r.t. the *left* eigenvectors of $\hat{\mathbf{A}}$, $g(Y) = \sum_k b_k u_k(Y)$. Substituting this into the definition of g^* yields after some calculation $g^*(s) = (1/\sqrt{|V|}) \sum_k b_k (\mu_k/D)^s$. ∎

5.5. Random Walks and Coherent Configurations

In this section we link the simple random walks on Γ with the properties of coherent configurations that are compatible with Γ. Our goal of course is to eventually derive a relation between the "random walk" correlation function $r(s)$ of a landscapes and its autocorrelation function ρ with respect to a coherent configuration $\mathfrak{R}$ on Γ. The first step is to consider the structure of $\mathbf{T}$ in some detail.

Lemma 11. Let Γ be a D-regular (connected) graph with adjacency matrix $\mathbf{A}$, and let $\mathbf{T} = (1/D)\mathbf{A}$ be the transition matrix of a simple random walk on V.
(a) If ϖ is an equitable partition of V anchored at $x_0 \in V$, then we have for each class $Y \in \varpi$ and all $y \in Y$:

$$[\mathbf{T}^s]_{yx_0} = \frac{1}{|Y|}\sum_{y\in Y}[\mathbf{T}^s]_{yx_0}.$$

(b) If $\mathfrak{R}$ is a *homogeneous* coherent configuration on V and compatible with Γ, then for each class $\mathcal{Y} \in \mathfrak{R}$ and each $x_0 \in V$ holds

$$\sum_{y\in Y}[\mathbf{T}^s]_{yx_0} = \frac{1}{|V|}\sum_{(y,x_0)\in\mathcal{Y}}[\mathbf{T}^s]_{yx_0} \stackrel{\text{def}}{=} \vartheta_{s\mathcal{Y}}\,. \tag{41}$$

We are now in the position to derive a simple geometric relationship between $r(s)$ and $\rho(\mathcal{X})$. The following lemma establishes a generalized version of a well known result for the distance classes of a distance transitive graph, see e.g. [15, 41].

Lemma 12. Let Γ be a connected D-regular graph, let $\mathfrak{R}$ be a homogeneous coherent configuration compatible with Γ and let f be a non-constant landscape on Γ, with "random walk" correlation function $r(s)$ and correlation function ρ with respect to $\mathfrak{R}$. Then

$$r(s) = \sum_{\mathcal{X}\in\mathfrak{R}}\vartheta_{s\mathcal{X}}\,\rho(\mathcal{X})\,. \tag{42}$$

Theorem 2. Let Γ be a connected D-regular graph with adjacency matrix $\mathbf{A}$ and let $\mathfrak{R}$ be a homogeneous coherent configuration on V compatible with Γ.

Furthermore let f be a non-flat landscape on Γ with "random walk" correlation function $r(s)$. Then:
The autocorrelation function ρ of f with respect to $\mathfrak{R}$ is a left eigenvector of $\hat{\mathbf{A}}$ if and only if $r(s)$ is an exponential.

Proof. Choose an arbitrary $x_0 \in V$ and consider the equitable partition $\mathfrak{R}_{x_0}$. Furthermore define $g(\mathcal{X}_{x_0}) \stackrel{\text{def}}{=} \rho(\mathcal{X})$ for all $\mathcal{X} \in \mathfrak{R}$. Then

$$g^*(s) \stackrel{\text{def}}{=} \sum_{\mathcal{Y}_{x_0} \in \mathfrak{R}_{x_0}} \sum_{y \in \mathcal{Y}_{x_0}} [\mathbf{T}^s]_{yx_0} g(\mathcal{Y}_{x_0}) = \sum_{\mathcal{Y} \in \mathfrak{R}} \vartheta_{s\mathcal{Y}} \rho(\mathcal{Y}) = r(s)$$

is the "random walk" correlation function of f as an immediate consequence of lemma 12 above. Proposition 14 guarantees that g^* is exponential if and only if g is a left eigenvector of $\hat{\mathbf{A}}$. Recalling that we may consider $\hat{\mathbf{A}}$ indexed by the classes of $\mathfrak{R}$ instead of by the classes of $\mathfrak{R}_{x_0}$ completes the proof. ∎

5.6. Association Schemes and Distance Regular Graphs

Suppose that $\mathfrak{R}$ is a symmetric coherent configuration (i.e., a (symmetric) association scheme) compatible with the D-regular graph Γ. Then the algebra $\langle \mathfrak{R} \rangle$ coincides with the adjacency algebra of Γ, i.e., any $\mathbf{X} \in \langle \mathfrak{R} \rangle$ is a polynomial in $\mathbf{A}$ [74]. Thus the ONB $\{\varphi_i\}$ of the Laplacian simultaneously diagonalizes all the matrices associated with the classes of $\mathfrak{R}$. The corresponding eigenvalues are known as the *eigenvalues of the association scheme*, $\mathbf{X}\varphi_i = \mathbf{p}_i(\mathcal{X})\varphi_i$ for all $\mathcal{X} \in \mathfrak{R}$, see [77, p.225].

Now consider the correlation function of a basis vector φ_i with respect to $\mathfrak{R}$:

$$\omega_i(\mathcal{X}) \stackrel{\text{def}}{=} \frac{|V|}{|\mathcal{X}|} \frac{\langle \varphi_i, \mathbf{X}\varphi_i \rangle}{\langle \varphi_i, \varphi_i \rangle} = \frac{|V|}{|\mathcal{X}|} \mathbf{p}_i(\mathcal{X}) . \tag{43}$$

As a consequence of theorems 1 and 2 we know that ω_i is a left eigenvector of the collapsed adjacency matrix, $\omega_i \hat{\mathbf{A}} = \mu_i \omega_i$. Proposition 11.iii implies that $v_i(\mathcal{X}) =$

$|\mathcal{X}|\omega_i(\mathcal{X})$ is the corresponding right eigenvector. Consequently we find that $\mathbf{p}_i$ itself is a right eigenvector of $\hat{\mathbf{A}}$. This is a well known result in the theory of distance regular graphs [82, sect. 4.1].

The functions ω_i fulfill the orthogonality relation

$$\sum_{\mathcal{X}\in\mathfrak{R}} |\mathcal{X}|\omega_i(\mathcal{X})\omega_j(\mathcal{X}) = \frac{|V|}{m(\lambda_i)}\,\delta_{i,j}\,, \tag{44}$$

where $m(\lambda_i)$ is the multiplicity of the i-th eigenvalue of the Laplacian of Γ, see [82, Prop. 2.2.2].

Even more is known if the graph Γ is distance regular. An association scheme is called P-polynomial if the matrices $\mathbf{D}_d$ corresponding to the distance classes can be written as polynomials of degree d in terms of the adjacency matrix $\mathbf{A}$. An association scheme is P-polynomial if and only if it consists of the distance classes of a distance regular graph, see e.g. [77, sect. 12.3]. The Hamming graphs $\mathcal{Q}^n_\alpha$ and the Johnson graphs $J(n,k)$ are of this type, while the Cayley graphs of the symmetric group are not distance regular, see [92] for the adjacency algebra of $\Gamma(S_4,\mathcal{T})$.

Suppose Γ is distance transitive; then for any two points $(x,y)\in V$ with distance $d(x,y) = d$ there are $c(d)$ neighbors of y with distance $d-1$ from x and $b(d)$ neighbors of y with distance $d+1$ from x. The functions $\omega_i(d) \stackrel{\text{def}}{=} \omega_i(\mathcal{D}_d)$ fulfill the following recursion:

$$\omega_i(d+1) = 1 + \frac{c(d)-\lambda_i}{b(d)}\omega_i(d) - \frac{c(d)}{b(d)}\omega_i(d-1) \qquad \omega_i(0)=1 \tag{45}$$

for $1\le d\le d^*$. In this expressions λ_i is as usual the eigenvalue of the graph Laplacian corresponding to ω_i. Furthermore it can be shown that the ω_i for a system of orthogonal polynomials [93]. We present here the explicit expressions for the Hamming and Johnson graphs, see also table 2.

Hamming graphs lead to Krawtchouk polynomials:

$$\omega_l(d) = \frac{1}{\binom{n}{l}(\alpha-1)^l}\mathbf{K}^{n,\alpha}_l(d) = \frac{1}{\binom{n}{l}(\alpha-1)^l}\sum_{j=0}^{l}(-1)^j(\alpha-1)^{l-j}\binom{d}{j}\binom{n-d}{l-j} \tag{46}$$

Table 2. Parameters for Hamming and Johnson graphs.

		$\mathcal{Q}_\alpha^n$	$J(n,k)$
Order	$\vert V\vert$	α^n	$\binom{n}{k}$
Degree	D	$(\alpha-1)n$	$k(n-k)$
Diameter	d^*	n	$\min[k,n-k]$
Eigenvalue	λ_j	$j\,\alpha$	$j(n-1+j)$
Multiplicity	$m(\lambda_j)$	$(\alpha-1)^j\binom{n}{j}$	$\frac{n+1-2j}{n+1-j}\binom{n}{j}$
Polynomial	ω_j	Krawtchouk, equ.(46)	Hahn, equ.(47)
	$b(d)$	$(\alpha-1)(n-d)$	$(k-d)(n-k-d)$
	$c(d)$	d	d^2

Krawtchouk polynomials play an important role in coding theory [94, 95].

The Hahn polynomials are the terminating hypergeometric series of type

$$ {}_3F_2(a_1,a_2,a_3;b_1,b_2;z) \stackrel{\text{def}}{=\!=} \sum_{j=0}^{\infty} \frac{(a_1)_j(a_2)_j(a_3)_j}{(b_1)_j(b_2)_j}\,\frac{z^j}{j!} $$

where Pochhammer's symbol $(a)_j$ is defined by $(a)_j = a(a+1)\dots(a+j-1)$ and $(a)_0 = 1$. The polynomials associated with the Johnson graphs are

$$ \begin{aligned} \omega_l^{n,k}(d) &= {}_3F_2(-d,-l,-n+(l-1);-(n-k),-k;1) \\ &= \sum_{j=0}^{l} \frac{(-l)_j\,(-n+(l-1))_j}{j!\,(k-n)_j\,(-k)_j}(-d)_j\,. \end{aligned} \tag{47} $$

6. Summary

A mathematical framework for the analysis of (fitness) landscapes on regular graphs has been developed. By a landscape we mean a given function f defined on finite set V of configurations together with a neighborhood relation between configurations that allows us to consider the "configuration space" as an undirected graph. The basic ingredient of this theory presented here is the "Fourier series" implied by the geometry of the graph that underlies the landscape. More precisely, we use an expansion of the landscape in terms of an orthonormal base of eigenfunctions of the Laplacian operator of the configuration space upon which the landscape f is built, see Sect 3.

We call a landscape elementary if it consists of an eigenfunction of the graph Laplacian plus an arbitrary constant function. The landscapes of a large number of important examples from spin glass physics to combinatorial optimization are of this type, among them the Sherrington-Kirkpatrick spin glass and the Travelling Salesman Problem, see Sect. 4.

Two types of correlation functions are commonly used for characterizing and comparing landscapes: "Time-series" sampled along random walks give rise to the "random walk" correlation function $r(s)$ of a landscape, see Sect. 2. Partitioning the set of all pairs of configurations into suitable classes, such as the classes induced by the natural distance measure between configurations. The autocorrelation function $\rho(\mathcal{X})$ of a given landscape with respect to a given partition of the configuration space can be defined in a canonical way. It has interesting properties provided the partition is consistent with symmetry properties of the configuration space. For example, the distance partition turns out to be the natural choice on sequence spaces. In more technical terms, the partition must form a coherent configuration compatible with the graph-structure of the configuration space, see Sect 5.

This contribution has been concerned with elucidating the relationships between the Fourier-expansion of a given landscapes and its two correlation measures $r(s)$ and $\rho(\mathcal{X})$. The main result of this paper characterizes elementary landscapes in terms of their autocorrelation functions. Summarizing theorems 1 and 2 we have:

Main Result. Let Γ be connected D-regular graph, and let $\mathfrak{R}$ be a homogeneous coherent configuration which is compatible with Γ. Furthermore let f be a non-flat landscape on Γ with "random walk" autocorrelation function r and autocorrelation function ρ with respect to $\mathfrak{R}$. Then the following statements are equivalent:

(i) f is elementary.

(ii) $f - \overline{f}\mathbf{1}$ is an eigenvector the Laplacian $-\Delta$ of Γ with a positive eigenvalue.

(iii) r is of the form $r(s) = a^{-s}$ (with $-1 \le a < 1$).

(iv) ρ is a left eigenvector of the collapsed adjacency matrix $\hat{\mathbf{A}}$ of Γ (with an eigenvalue smaller than D).

Since any landscape on Γ is necessarily a superposition of elementary landscapes, so are their correlation functions. Thus suppose $f = \sum_j a_j \varphi_j$, where $-\Delta\varphi_k = \lambda_k \varphi_k$. Let λ_p, $p = 0, \dots, M$ designate only the distinct eigenvalues, and let I_p be an index set such that $\lambda_k = \lambda_p$ if and only if $k \in I_p$. As usual we use $\lambda_0 = 0$ for the smallest eigenvalue of $-\Delta$. Define

$$B_p \stackrel{\text{def}}{=\!=} \sum_{k \in I_p} A_p = \frac{\sum_{k \in I_p} |a_k|^2}{\sum_{k \neq 0} |a_k|^2}$$

for all $p \neq 0$. The we have

$$r(s) = \sum_{p \neq 0} B_p (1 - \lambda_p / D)^s .$$

If furthermore all eigenvalues of the collapsed adjacency matrix $\hat{\mathbf{A}}$ for which there are non-zero coefficients in the Fourier expansion of f are simple then we have

$$\rho(\mathcal{X}) = \sum_{p \neq 0} B_p \frac{u_p(\mathcal{X})}{u_p(\mathcal{I})},$$

where u_p is a left eigenvector of $\hat{\mathbf{A}}$ belonging to the eigenvalue λ_p. This is true in particular for all landscapes on Hamming graphs and Johnson graphs.

7. Discussion

Many important examples of landscapes are elementary, i.e., (up to an additive constant) they fulfill a discrete analogue of the Helmholtz equation $\Delta f + \lambda f = 0$. Among them are Derrida's p-spin models, and the landscapes of the best known combinatorial optimization problems. Elementary landscapes exhibit a characteristic distribution of local optima on the configuration space which depends crucially on the corresponding eigenvalue λ of the Laplacian. In particular, the "location" of λ in the spectrum of the Laplace operator determines the maximum number of nodal domains, that is the maximum number of disconnected islands of values of f that are above average. The analogy with the properties of the eigenfunctions of the Laplacian on a Riemannian manifolds strongly suggests the conjecture that a landscapes has to be more rugged if it corresponds to a higher "excited state" in the spectrum of the Laplacian. In fact, the nearest neighbor correlation on the landscape is directly linked to the eigenvalue λ. More precisely, the set of all autocorrelation functions forms a simplex spanned by suitably normalized left eigenvectors of the collapsed adjacency matrix of the configuration space. The corresponding eigenvalues determine the decay of the correlation function along a simple random walk on the landscape.

The relation between the Fourier expansion of a landscape and the computational complexity of the optimization problem on the landscape is of great importance for devising practical optimization heuristics. Disappointingly, there seems to be **no** simple relation. Most of the elementary landscapes discussed in this paper, such as the `TSP`, are $\mathcal{NP}$-complete, but have only a fairly small number of non-zero Fourier coefficients, namely those corresponding to a single eigenvalue of the configuration space.

The formalism derived in this contribution suggests to *approximate* a given landscape by a superposition of elementary landscapes, in particular if we do not have a closed form but only a computationally costly algorithm for evaluating it at particular configurations such as, for example, in the case of RNA secondary structure models. The elementary landscapes depend only on the configuration space and can be obtained explicitly in many cases, in particular for all sequences spaces with constant chain length (i.e., Hamming graphs). In practice it is even easier to directly compute the correlation functions of elementary landscapes. By the Main Theorem one needs only the collapsed adjacency matrix of the configuration space, which is small enough in many cases that numerical solutions can be obtains even if an analytical expression is unknown, as in the case of TSPs. A comparison of the autocorrelation functions obtained from computational studies of the landscape of interest on the correlation functions of elementary landscapes on the same configuration space allow to estimate the *amplitudes* B_i, that is the coefficients of the decomposition of the "experimental" correlation function into the elementary ones. These amplitudes can be used as an easy means of describing the landscape. On a Boolean hypercube, for instance, the amplitudes B_p have a very intuitive interpretation: they measure the relative importance of p-ary (spin) interactions. As an example we have seen in sect. 4.2 that the landscape of the "Low Autocorrelated Binary String Problem" consists of two "modes": An asymptotically vanishing contribution of a 2-spin model and a dominating mode corresponding to a 4-spin model. Asymmetric TSPs, as another example, consists of two modes, corresponding to the symmetric and the anti-symmetric part of the cost matrix. A study on the decomposition of RNA landscapes into "modes" corresponding to elementary landscapes will be reported elsewhere [96]. Of course, even a complete understanding of elementary landscapes is only the first chapter in the story of landscapes: What can be said about the superposition of elementary landscapes, even if the structure of the elementary parts is known in detail?

Many open questions remain. Are there better bounds than Courant's theorem on the number of nodal domain of an elementary landscape? What is the relation between eigenvalue and number of local optima in an elementary landscape? How much of the theory outlined in this contribution carries over to configuration space with less symmetry, such as spaces of finite trees? Can we devise a comparable formalism for the combinatory maps of sequence-structure relations [16], where the image of a configuration is not a real-valued fitness but an element of an abstract metric space?

Acknowledgments

Stimulating discussions with P. Anderson, B. Davies, A.W.M. Dress, W. Fontana, R. Palmer, P. Phillipson, Chr. Reidys, and G.P. Wagner are gratefully acknowledged. The author is grateful to the German *Max Planck Gesellschaft* for supporting his stay at the Santa Fe Institute in 1993 where the work on this manuscript began.

References

[1] S. Wright. The roles of mutation, inbreeding, crossbreeeding and selection in evolution. In D. F. Jones, editor, *Int. Proceedings of the Sixth International Congress on Genetics*, volume 1, pages 356–366, 1932.

[2] K. Binder and A. P. Young. Spin glasses: Experimental facts, theoretical concepts, and open questions. *Rev.Mod.Phys.*, 58:801–976, 1986.

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

[4] M. Garey and D. Johnson. *Computers and Intractability. A Guide to the Theory of $\mathcal{NP}$ Completeness.* Freeman, San Francisco, 1979.

[5] M. Eigen. Selforganization of matter and the evolution of biological macromolecules. *Die Naturwissenschaften*, 10:465–523, 1971.

[6] W. Fontana and P. Schuster. A computer model of evolutionary optimization. *Biophysical Chemistry*, 26:123–147, 1987.

[7] W. Fontana, W. Schnabl, and P. Schuster. Physical aspects of evolutionary optimization and adaption. *Physical Review A*, 40(6):3301–3321, 1989.

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

[9] E. L. Lawler, J. K. Lenstra, A. H. G. R. Kan, and D. B. Shmoys. *The Traveling Salesman Problem. A Guided Tour of Combinatorial Optimization.* John Wiley & Sons, 1985.

[10] Y. Fu and P. W. Anderson. Application of statistical mechanics to NP-complete problems in combinatorial optimization. *J.Phys.A:Math.Gen.*, 19:1605–1620, 1986.

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

[12] 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.

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

[14] S. Bonhoeffer, J. McCaskill, P. Stadler, and P. Schuster. Temperature dependent RNA landscapes, a study based on partition functions. *European Biophysics Journal*, 22, 1993.

[15] W. Fontana, T. Griesmacher, W. Schnabl, P. Stadler, and P. Schuster. Statistics of landscapes based on free energies, replication and degredation rate constants of RNA secondary structures. *Monatsh. Chemie*, 122:795–819, 1991.

[16] 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(3):2083 – 2099, 1993.

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

[18] 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.

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

[20] P. Schuster. Complex optimization in an artificial RNA world. In D. Farmer, C. Langton, S. Rasmussen, and C. Taylor, editors, *Artificial Life II*, volume XII of *SFI Studies in the Science of Complexity*, pages 277–291. Addison-Wesley, 1992.

[21] 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.

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

[23] P. F. Stadler and W. Grüner. Anisotropy in fitness landscapes. *J. Theor. Biol.*, 165:373–388, 1993.

[24] M. Tacker, W. Fontana, P. Stadler, and P. Schuster. Statistics of RNA melting kinetics. *Eur. J. Biophys.*, 23:29–38, 1994.

[25] C. Macken, P. Hagan, and A. Perelson. Evolutionary walks on rugged landscapes. *SIAM J.Appl.Math.*, 51:799–827, 1991.

[26] H. Flyvbjerg and B. Lautrup. Evolution in a rugged fitness landscape. *Phys. Rev. A*, 46:6714–6723, 1992.

[27] P. Bak, H. Flyvbjerg, and B. Lautrup. Coevolution in a rugged fitness landscape. *Phys.Rev. A[15]*, 46:6724–6730, 1992.

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

[29] S. A. Kauffman and E. D. Weinberger. The n-k model of rugged fitness landscapes and its application to maturation of the immune response. *J. Theor. Biol.*, 141:211, 1989.

[30] E. D. Weinberger. Local properties of Kauffman's N-k model: A tunably rugged energy landscape. *Phys. Rev. A*, 44(10):6399–6413, 1991.

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

[32] B. Derrida. Random-energy model: An exactly solvable model of disorderes systems. *Phys. Rev. B*, 24(5):2613–2626, 1981.

[33] C. Amitrano, L. Peliti, and M. Saber. Population dynamics in a spin-glass model of chemical evolution. *J. Mol. Evol.*, 29:513–525, 1989.

[34] E. Weinberger and P. Stadler. Why some fitness landscapes are fractal. *J. Theor. Biol.*, 163:255–275, 1993.

[35] A. Bray and M. Moore. Metastable states in spin glasses. *J.Phys.C:Solid St.Phys.*, 13:L469–L476, 1980.

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

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

[38] M. Zuker and D. Sankoff. RNA secondary structures and their prediction. *Bull.Math.Biol.*, 46(4):591–621, 1984.

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

[40] J. Maynard-Smith. Natural selection and the concept of a protein space. *Nature*, 225:563–564, 1970.

[41] E. D. Weinberger. Correlated and uncorrelated fitness landscapes and how to tell the difference. *Biol.Cybern.*, 63:325–336, 1990.

[42] F. Spitzer. *Principles of Random Walks.* Springer-Verlag, New York, 1976.

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

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

[45] 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.

[46] P. M. Soardi. *Potential Theory on Infinite Networks*, volume 1590 of *Lecture Notes in Mathematics.* Springer-Verlag, Berlin, 1994.

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

[48] B. Huppert. *Endliche Gruppen I.* Springer-Verlag, Berlin, Heidelberg, 1967.

[49] J.-P. Serre. *Linear Representations of Finite Groups.* Springer-Verlag, New York, Heidelberg, Berlin, 1977.

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

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

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

[53] B. Derrida. The random energy model. *Phys.Rep.*, 67:29–35, 1980.

[54] E. Gardner and B. Derrida. The probability distribution of the partition function of the random energy model. *J.Phys. A*, 22:1975–1982, 1989.

[55] M. Golay. Sieves for low-autocorrelation binary sequences. *IEEE Trans.Inform.Th.*, IT-23:43–51, 1977.

[56] J. Bernasconi. Low autocorrelation binary sequences: statistical mechanics and configuration space analysis. *J.Physique*, 48:559–567, 1987.

[57] D. Rumschitzky. Spectral properties of Eigen's evolution matrices. *J.Math.Biol.*, 24:667–680, 1987.

[58] A. Dress and D. Rumschitzki. Evolution on sequence space and tensor products of representation spaces. *Acta Appl.Math.*, 11:103–111, 1988.

[59] S. Lin and B. Kernighan. An effective heuristic algorithm for the traveling salesman problem. *Oper.Res.*, 21:498–516, 1965.

[60] J. Warren. Special cases of the travelling salesman problem. *Appl.Math.& Comp.*, 60:171–177, 1994.

[61] B. Krakhofer. Local optima in landscapes of combinatorial optimization problems. Master's thesis, University of Vienna, Dept. of Theoretical Chemistry, 1995.

[62] E. Aarts and J. Korst. *Simulated Annealing and Boltzman Machines.* J. Wiley & Sons, New York, 1990.

[63] R. Otten and L. vanGinneken. *The Annealing Algorith.* Kluwer Acad.Publ., Boston, 1989.

[64] D. Miller and J. F. Pekny. Exact solution of large asymmetric traveling salesman problems. *Science*, 251:754–761, 1991.

[65] B. Mohar. The laplacian spectrum of graphs. In Y. Alavi, G. Chartrand, O. Ollermann, and A. Schwenk, editors, *Graph Theory, Combinatorics, and Applications*, pages 871–897, New York, 1991. John Wiley & Sons.

[66] I. Chavel. *Eigenvalues in Riemannian Geometry.* Academic Press, Orlando Fl., 1984.

[67] Y. Colin De Verdiere. Multiplicites des valeurs prores laplaciens discrets et laplaciens continus. *Rend. mat. appl.*, 13:433–460, 1993.

[68] M. Fiedler. Algebraic connectivity of graphs. *Czechoslovak Math. J.*, 23:298–305, 1973.

[69] R. Grone and G. Zimmermann. Large eigenvalues of the laplacian. *Lin. & Multilin. Algebra*, 28:45–47, 1990.

[70] R. Merris. Characteristic vertices of trees. *Lin. & Multilin. Algebra*, 22:115–131, 1987.

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

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

[73] F. Buckley and F. Harrary. *Distance in Graphs.* Addison Wesley, Redwood City CA, 1990.

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

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

[76] D. Higman. Coherent algebras. *Lin. Alg. Appl.*, 93:209–239, 1987.

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

[78] P. Cameron and J. vanLint. *Designs, Graphs, Codes, and their Links*, volume 22 of *London Math. Soc. Student Texts.* Cambridge University Press, Cambridge UK, 1991.

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

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

[81] P. F. Stadler and R. Happel. Canonical approximation of landscapes. Santa Fe Institute Preprint 94-09-51, 1994.

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

[83] N. Biggs. *Finite Groups of Automorphisms*, volume 6 of *London Mathematical Society Lecture Notes.* Cambridge University Press, Cambridge UK, 1971.

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

[85] G. Sabidussi. Vertex transitive graphs. *Mh. Math*, 68:426–438, 1964.

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

[87] 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.

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

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

[90] P. F. Stadler. Linear operators on correlated landscapes. *J.Physique*, 4:681–696, 1994.

[91] P. F. Stadler. Random walks and orthogonal functions associated with highly symmetric graphs. *Disc. Math.*, 1995. in press, Santa Fe Institute Preprint 93-06-31.

[92] P. F. Stadler. Towards a theory of landscapes. In R. López Peña, editor, *Complex Systems and Binary Networks*. Springer-Verlag, New York, 1995. in press, Santa Fe Institute Preprint 95-03-30.

[93] N. Sloane. An introduction to association schemes and coding theory. In R. A. Askey, editor, *Theory and Applications of Special Functions*, pages 225–260. Academic Press, New York, San Franscisco, London, 1975.

[94] S. Roman. *Coding and Information Theory*. Springer-Verlag, New York, Heidelberg, 1992.

[95] J. vanLint. *Introduction to Coding Theory*. Springer-Verlag, New York, 1982.

[96] R. Happel and P. F. Stadler. Canonical approximation of rna landscapes. in preparation, 1995.

Appendix: Proofs

Proof of Lemma 1.

Let $\delta_{x_0}(x) = \delta_{x,x_0}$. Then the probability that $x_t = z$ is given by the z-coordinate of $[\mathbf{T}^s\delta_{x_0}]$. Thus

$$\begin{aligned}
\langle F(x_t)\rangle_{x_0,t} &= \lim_{T\to\infty}\frac{1}{T+1}\sum_{t=0}^{T}\frac{1}{|V|}\sum_{x_0\in V}\sum_{z\in V}F(z)[\mathbf{T}^t\delta_{x_0}]_z \\
&= \lim_{T\to\infty}\frac{1}{T+1}\sum_{t=0}^{T}\frac{1}{|V|}\sum_{z\in V}F(z)\left[\mathbf{T}^t\left(\sum_{x_0\in V}\delta_{x_0}\right)\right]_z \\
&= \lim_{T\to\infty}\frac{1}{T+1}\sum_{t=0}^{T}\frac{1}{|V|}\sum_{z\in V}F(z)[\mathbf{T}^t\mathbf{1}]_z
\end{aligned}$$

Since Γ is regular we have $\mathbf{T1}=\mathbf{1}$, and therefore

$$\langle F(x_t)\rangle_{x_0,t} = \lim_{T\to\infty}\frac{1}{T+1}\sum_{t=0}^{T}\frac{1}{|V|}\sum_{z\in V}F(z) = \lim_{T\to\infty}\frac{1}{T+1}\sum_{t=0}^{T}\overline{F}.$$

Since $\overline{F}$ is constant in "time" we have $\langle F(x_t)\rangle_{x_0,t}=\overline{F}$. ∎

Proof of Lemma 2.

As in the proof of lemma 1 we start with

$$\begin{aligned}
\langle F(x_{t+s},x_t)\rangle_{x_0,t} &= \lim_{T\to\infty}\frac{1}{T+1}\sum_{t=0}^{T}\frac{1}{|V|}\sum_{x_0\in V}\sum_{z,z'\in V}F(z',z)[\mathbf{T}^s\delta_z]_{z'}[\mathbf{T}^t\delta_{x_0}]_z \\
&= \lim_{T\to\infty}\frac{1}{T+1}\sum_{t=0}^{T}\frac{1}{|V|}\sum_{z,z'\in V}F(z',z)[\mathbf{T}^s\delta_z]_{z'}\left[\mathbf{T}^t\left(\sum_{x_0\in V}\delta_{x_0}\right)\right]_z \\
&= \lim_{T\to\infty}\frac{1}{T+1}\sum_{t=0}^{T}\frac{1}{|V|}\sum_{z,z'\in V}F(z',z)[\mathbf{T}^s\delta_z]_{z'}\cdot 1 \\
&= \frac{1}{|V|}\sum_{z,z'\in V}F(z',z)[\mathbf{T}^s\delta_z]_{z'}
\end{aligned}$$

By definition we have $[\mathbf{T}^s\delta_z]_{z'}=[\mathbf{T}^s]_{z,z'}$; this completes the proof. ∎

Proof of Corollary 1.

Use lemma 1 with $F=f$ and $F=f^2$, respectively, and lemma 2 with $F(x,y)=f(x)f(y)$. Substituting the results into the definition of $r(s)$ completes the proof.

∎

Proof of Lemma 3.

$\Delta\mathbf{1} = 0$ implies that $\tilde{f} = f - f^*\mathbf{1}$ is an eigenvector of $-\Delta$ with eigenvalue λ. Thus $\tilde{f}$ either constant and $\lambda = 0$, or $\lambda > 0$ and $\langle \tilde{f}, \mathbf{1}\rangle = 0$. In the first case f itself is a flat landscape with $f(x) = f^* = \overline{f}$ for all $x \in V$. In the second case we have

$$|V|\overline{f} = \sum_{x\in V} f(x) = \langle \tilde{f}, \mathbf{1}\rangle + f^*\langle \mathbf{1}, \mathbf{1}\rangle = 0 + f^*|V|\,.$$

Consequently, if f is elementary, then the constant f^* coincides with the mean value $\overline{f}$ of the landscape f. Connectedness of Γ implies that 0 is a single eigenvalue and thus $\tilde{f}$ must be an eigenvector belonging to an eigenvalue $\lambda > 0$.
The converse is trivial: a landscape of the form (6) is always elementary. ∎

Proof of Lemma 4.

The Laplacian $-\Delta$ of K_n has only two distinct eigenvalues, $\lambda_0 = 0$ with multiplicity 1 and eigenvector $\mathbf{1}$, and the $n-1$-fold degenerate eigenvalue $\lambda_1 = n$ [51]. Thus any non-constant f is of the form $c\mathbf{1} + \varphi$ with some constant $c \in R$ and $\langle \varphi, \mathbf{1}\rangle = 0$, i.e., φ is an eigenvector belonging to λ_1. ∎

Proof of Lemma 5.

It is convenient to allow arbitrary indices, setting $\sigma_q = 0$ for all $q < 1$ or $q > n$. With this convention we have

$$\begin{aligned}
f(\sigma) &= \sum_{k=1}^{n-1}\sum_{i=1}^{n-k}\sum_{j=1}^{n-k} \sigma_i\sigma_{i+k}\sigma_j\sigma_{j+k} \\
&= \sum_{k=1}^{n-1}\sum_{i=1}^{n-k}\left[\sum_{j\neq i-k,i,i+k} \sigma_i\sigma_{i+k}\sigma_j\sigma_{j+k} + \sigma_{i-k}\sigma_i^2\sigma_{i+k} + \sigma_i^2\sigma_{i+k}^2 + \sigma_i\sigma_{i+k}^2\sigma_{i+2k}\right] \\
&= \frac{n(n-1)}{2} + \sum_{k=1}^{n-1}\sum_{i=1}^{n-k}\left[\sigma_{i-k}\sigma_{i+k} + \sigma_i\sigma_{i+2k}\right] + \sum_{k=1}^{n-1}\sum_{i=1}^{n-k}\sum_{j\neq i-k,i,i+k} \sigma_i\sigma_{i+k}\sigma_j\sigma_{j+k}
\end{aligned}$$

Using the definition of $\varepsilon_{\dots}(\sigma)$ and eliminating all terms containing a σ_q with $q < 1$ or $q > n$ completes the proof. ∎

Proof of Corollary 2.

The functions $\varepsilon_{i_1,i_2,\ldots,i_p}$, where all indices i_k are different, are eigenfunctions of the Laplacian $-\Delta$ of the Boolean hypercube belonging to the eigenvalues λ_p. Equ.(9) in the proof of theorem 1 implies immediately that $r(s) = \frac{a}{a+b}(1-4/n)^s + \frac{b}{a+b}(1-8/n)^s$ where a and b are the sums of the squares of the coefficients of ε_2 and ε_4, respectively. Since all coefficients are 0 or 1 we simply need to count all non-zero terms in lemma 5. We find that $a = \mathcal{O}(n^2)$ and $b = \mathcal{O}(n^3)$, and the corollary follows. ∎

Proof of Lemma 6.

Without loosing generality we can consider the neighborhood of the identity permutation $\imath$, since the numbering of the cities is arbitrary.

(i) In the case of transpositions we start with Grover's [47] formula

$$\begin{aligned} f((i,j)) - f(\imath) = \quad & w_{j,i-1} + w_{i+1,j} + w_{i,j-1} + w_{j+1,i} \\ & - w_{i,i-1} - w_{i+1,i} - w_{j,j-1} - w_{j+1,j} \\ & + (\delta_{j,i+1} + \delta_{j,i-1})(w_{ij} + w_{ji}). \end{aligned}$$

Summing over all $i \neq j$ yields $2\Delta f(\imath) = 4X - 4nf(\imath) + 2f(\imath) + 2f^*(\imath)$ and analogously we obtain $2\Delta f^*(\imath) = 4X - 4nf^*(\imath) + 2f^*(\imath) + 2f(\imath)$, where X is the sum over all non-diagonal entries of W. The proposition follows immediately.

(ii) In case of inversions we find

$$f([k,l]) - f(\imath) = w_{l,k-1} + w_{l+1,k} - w_{k,k-1} - w_{l+1,l} + \sum_{j=k}^{l-1}(w_{j,j+1} - w_{j+1,j})\,.$$

Summing over all $k \neq l$ yields $2\Delta f(\imath) = 2X - 2nf(\imath) + \dfrac{n(n-1)}{2}[f^*(\imath) - f(\imath)]$. A completely analogous result is obtained for $\Delta f^*(\imath)$, and a short calculation then completes the proof. ∎

Proof of Lemma 7.

Without loosing generality we may evaluate the cost function of the identity permutation $\imath$ since the labeling of the vertices is arbitrary. Using the labels k of the matched pairs as indices we have to following five types of neighbors, and the corresponding changes of the cost function for each of them

[11'] $[W_{2j-1,2i} + W_{2i-1,2j}] - [W_{2i-1,2i} + W_{2j-1,2j}]$,
[12'] $[W_{2j,2i} + W_{2j-1,2i-1}] - [W_{2i-1,2i} + W_{2j-1,2j}]$,
[21'] $[W_{2i-1,2j-1} + W_{2i,2j}] - [W_{2i-1,2i} + W_{2j-1,2j}]$,
[22'] $[W_{2i-1,2j} + W_{2j-1,2i}] - [W_{2i-1,2i} + W_{2j-1,2j}]$, and
[11] $W_{2i,2i-1} - W_{2i-1,2i}$.

Summing the first four terms over all $i \neq j$ (i.e., counting each neighbor twice) yields

$$\begin{aligned} &[oe] + [oe] - 2f(\imath) - 2(n/2-1)f(\imath) \quad + \quad [ee] + [oo] - d - 2(n/2-1)f(\imath) \quad + \\ &[oo] + [ee] - d - 2(n/2-1)f(\imath) \quad + \quad [oe] + [oe] - 2f(\imath) - 2(n/2-1)f(\imath), \end{aligned}$$

where $[oe]$, etc., denotes the sum over all W_{ij} with odd i and even j, and d is the sum over all diagonal entries W_{kk}. Using that W is symmetric, i.e., $[oe] = [eo]$, this sum my be written as $2\Delta f(\imath) = 2\sum_{i\neq j} W_{ij} - 4(n-1)f(\imath)$, since the transposition $(2i-1, 2i)$ gives a contribution of 0. Thus we have $\Delta f(\imath) = 2(n-1)\,[\overline{f} - f(\imath)]$, i.e., $\lambda = 2(n-1)$, and f is elementary on the Cayley graph $\Gamma(S_n, \mathcal{T})$. ∎

Proof of Lemma 8.

From the definition of ρ we find

$$\sum_{\mathcal{X}\in\mathfrak{R}} \rho(\mathcal{X}) \cdot |\mathcal{X}| = \frac{|V|}{\langle f, f\rangle} \left\langle f, \left(\sum_{\mathcal{X}\in\mathfrak{R}} \mathbf{X}\right) f \right\rangle = \frac{|V|}{\langle f, f\rangle} \langle f, \mathbf{J}f\rangle = \frac{|V|}{\langle f, f\rangle}\,.$$

where $\mathbf{J}$ is the matrix with all entries 1. It is straightforward to check $\mathbf{J}f = |V|\overline{f}\mathbf{1}$, and thus $\langle f, \mathbf{J}f\rangle = |V|^2\overline{f}^2 = 0$. ∎

Proof of Lemma 9.

This statement is equivalent to claiming that $\mathcal{X} \cap \mathcal{D}_d \neq \emptyset$ implies $\mathcal{X} \subseteq \mathcal{D}_d$ if and only if $\mathfrak{R}$ is compatible with Γ. The "only if" part is trivial.

Assume thus that $\mathfrak{R}$ is compatible with Γ. We proceed by induction in d. The claim is true for $d = 0$ since $\mathfrak{R}$ is pre-coherent, and for $d = 1$ by compatibility with Γ. Suppose now, that the claim is true for all distances up to d and consider a pair of vertices (x, y) with distance $d(x, y) = d + 1$. Let $\mathcal{Z}$ be the class of $\mathfrak{R}$ to which (x, y) belongs. The triangle inequality implies that there is $z \in V$ such that $d(x, z) = d$ and $d(z, y) = 1$ and let $(x, z) \in \mathcal{X}$ and $(z, y) \in \mathcal{Y} \subseteq \mathcal{A}$. Therefore $p^{\mathcal{Z}}_{\mathcal{X}\mathcal{Y}} > 0$. Now consider an arbitrary pair $(x', y') \in \mathcal{Z}$. Since $p^{\mathcal{Z}}_{\mathcal{X}\mathcal{Y}} > 0$ there is at least one $z' \in V$ which fulfills $d(x', z') = d$ and $d(z', y') = 1$. The triangle inequality implies $d(x', y') \leq d + 1$. If $d(x', y') < d + 1$ we have $\mathcal{Z} \cap \mathcal{D}_h \neq \emptyset$ for some $h \leq d$ and thus, by the induction hypothesis, $\mathcal{Z} \subseteq \mathcal{D}_h$. This contradicts $d(x, y) = d + 1$. We conclude $d(x', y') = d + 1$ and therefore $\mathcal{Z} \subseteq \mathcal{D}_{d+1}$. ∎

Proof of Lemma 10.

Consider three classes $\mathcal{X}, \mathcal{Y}, \mathcal{Z} \in \mathfrak{R}$ and an arbitrary vertex $x_0 \in V$. We have

$$\begin{aligned} p^{\mathcal{Z}}_{\mathcal{X}\mathcal{Y}} &= \left|\left\{z \in V \middle| (x, z) \in \mathcal{X} \text{ and } (z, y) \in \mathcal{Y}\right\}\right| \quad \forall\, (x, x_0) \in \mathcal{Z} \\ &= \left|\left\{z \in \mathcal{Y}_{x_0} \middle| (x, z) \in \mathcal{X}\right\}\right| \quad \forall\, x \in \mathcal{Z}_{x_0} \end{aligned}$$

On the other hand we obtain

$$\hat{\mathbf{A}}_{\mathcal{X}_{x_0}\mathcal{Y}_{x_0}} = \left|\left\{z \in \mathcal{Y}_{x_0} \middle| (x, z) \in \mathcal{A}\right\}\right| = \sum_{\mathcal{X} \subseteq \mathcal{A}} \left|\left\{z \in \mathcal{Y}_{x_0} \middle| (x, z) \in \mathcal{X}\right\}\right|$$

for all $x \in \mathcal{Z}_{x_0}$. Comparing this with the above representation of the intersection numbers yields

$$\hat{\mathbf{A}}_{\mathcal{X}_{x_0}\mathcal{Y}_{x_0}} = \sum_{\substack{\mathcal{Z} \in \mathfrak{A} \\ \mathcal{Z} \cap \mathcal{A} \neq \emptyset}} p^{\mathcal{Z}}_{\mathcal{X}\mathcal{Y}} \,,$$

which is independent of the representatives of the classes by definition. Thus $\mathfrak{R}_{x_0}$ is equitable. It remains to show that $\mathfrak{R}_{x_0}$ is in fact anchored in x_0. $\mathcal{X} \subseteq \mathcal{I}$ implies that $\mathcal{X}_{x_0}$ is either empty $\mathcal{X}_{x_0} = \{x_0\}$. By definition there is some $\mathcal{X} \subseteq \mathcal{I}$ that contains (x_0, x_0), and thus $\{x_0\} \in \mathfrak{R}_{x_0}$. ∎

Proof of Lemma 11.

(a) By proposition 11.ix we know that $[\mathbf{A}^s]_{yx_0} = [\mathbf{A}^s]_{y'x_0}$ whenever y and y' are in the same class of ϖ, and thus the same holds true for the powers of $\mathbf{T}$. Thus we have for all $y \in Y$

$$\sum_{y\in Y} [\mathbf{T}^s]_{yx_0} = |Y|\,[\mathbf{T}^s]_{yx_0}\,.$$

(b) Since $\mathfrak{R}$ is a coherent configuration and compatible with Γ we may write $\mathbf{A}^s = \sum_{\mathcal{X}\in\mathfrak{R}} b_{\mathcal{X}}\mathbf{X}$, and thus $\mathbf{A}^s$ is constant on the classes of $\mathfrak{R}$. Furthermore, we have shown in lemma 10 that the projection $\mathcal{Y}_{x_0}$ is an equitable partition anchored at x_0. Thus we have $\sum\limits_{y\in\mathcal{Y}_{x_0}} [\mathbf{T}^s]_{yx_0} = |\mathcal{Y}_{x_0}|\,[\mathbf{T}^s]_{yx_0}$, where the second factor on the r.h.s. depends only on the class $\mathcal{Y}$ to which the pair (y, x_0) belongs. The first factor is

$$|\mathcal{Y}_{x_0}| = \left\{y\in V \,\middle|\, (y,x_0)\in\mathcal{Y}\right\} = \left\{y\in V \,\middle|\, (y,x_0)\in\mathcal{Y} \text{ and } (x_0,y)\in\mathcal{Y}^+\right\} = p^{\mathcal{X}}_{\mathcal{Y}^+\mathcal{Y}}\,,$$

where $\mathcal{X}$ is the class of $\mathfrak{R}$ to which (x_0, x_0) belongs. Since $\mathfrak{R}$ is homogeneous by assumption we have $\mathcal{X} = \mathcal{I}$, and thus $|\mathcal{Y}_{x_0}| = p^{\mathcal{I}}_{\mathcal{Y}^+\mathcal{Y}}$ for all $x_0 \in V$. Consequently we have $|\mathcal{Y}| = |V|\cdot|\mathcal{Y}_{x_0}|$ and the first factor is independent of x_0 as well. We have

$$\sum_{(y,x_0)\in\mathcal{Y}} [\mathbf{T}^s]_{yx_0} = \sum_{x_0\in V}\sum_{y\in\mathcal{Y}_{x_0}} [\mathbf{T}^s]_{yx_0}$$

and part (b) of the lemma follows immediately. ∎

Proof of Lemma 12.

We begin with the definition of $r(s)$. Without loosing generality we assume $\overline{f} = 0$ and $\sigma_f^2 = 1$. Thus

$$r(s) = \frac{1}{|V|}\sum_{x_0\in V}\sum_{y\in V}[\mathbf{T}^s]_{yx_0} f(y)f(x_0) = \frac{1}{|V|}\sum_{x_0\in V}\sum_{\mathcal{Y}_{x_0}\in\mathfrak{R}_0}\sum_{y\in\mathcal{Y}_{x_0}}[\mathbf{T}^s]_{yx_0} f(y)f(x_0)\,.$$

As a consequence of lemma 11.a we have

$$r(s) = \frac{1}{|V|}\sum_{x_0\in V}\sum_{\mathcal{Y}_{x_0}\in\mathfrak{R}_0}\left(\sum_{z\in\mathcal{Y}_{x_0}}[\mathbf{T}^s]_{yx_0}\right)\times\frac{1}{|\mathcal{Y}_{x_0}|}\sum_{y\in\mathcal{Y}_{x_0}} f(y)f(x_0)\,,$$

and using lemma 11.b we may rearrange this as

$$\begin{aligned} r(s) &= \frac{1}{|V|} \sum_{x_0 \in V} \sum_{\mathcal{Y}_{x_0} \in \mathfrak{R}_0} \vartheta_{s\mathcal{Y}} \frac{1}{|\mathcal{Y}_{x_0}|} \sum_{y \in \mathcal{Y}_{x_0}} f(y) f(x_0) \\ &= \sum_{\mathcal{Y} \in \mathfrak{R}} \vartheta_{s\mathcal{Y}} \frac{1}{|V|} \sum_{x_0 \in V} \frac{1}{|\mathcal{Y}_{x_0}|} \sum_{y \in \mathcal{Y}_{x_0}} f(y) f(x_0) = \sum_{\mathcal{Y} \in \mathfrak{R}} \vartheta_{s\mathcal{Y}} \frac{1}{|\mathcal{Y}|} \sum_{(x,y) \in \mathcal{Y}} f(x) f(y) \,. \end{aligned}$$

Substituting the definition of $\rho(\mathcal{Y})$ completes the proof. ∎