

# Average Fitness Differences on NK Landscapes

Wim Hordijk<sup>\*,a</sup> · Stuart A. Kauffman<sup>b</sup> · Peter F. Stadler<sup>c-x</sup>

*In memoriam Manfred Eigen \*9 May 1927 – † 6 Feb 2019*

Manuscript date Sun February 10 2019

**Abstract** The average fitness difference between adjacent sites in a fitness landscape is an important descriptor that impacts in particular the dynamics of selection/mutation processes on the landscape. Of particular interest is its connection to the error threshold phenomenon. We show here that this parameter is intimately tied to the ruggedness through the landscape’s amplitude spectrum. For the NK model a surprisingly simple analytical estimate explains simulation data with high precision.

**Keywords** NK model · Fitness landscapes · Elementary Landscape · Graph Laplacian · Error threshold

## 1 Introduction

Fitness landscapes have been studied for nearly a century as a convenient model of evolutionary adaptation. In this picture, fitness takes the role of a potential function with the combined effects of mutation and selection driving a popula-

tion uphill (Wright, 1932, 1967). The dynamics of this process crucially depends on the “topography” of the landscape, which in turn can be understood as a combination of a topological structure on the underlying search space  $X$  that induces a notion of adjacency or nearness (Stadler and Stadler, 2002), and properties of the fitness function  $f : X \rightarrow \mathbb{R}$  itself. Due to the generality of the model, fitness landscapes have been investigated in many different contexts beyond evolutionary biology, including spin glass physics, biopolymer folding, metabolic adaptation, gene regulation, and combinatorial optimization, see e.g. (Reidys and Stadler, 2002; Richter and Engelbrecht, 2014). Recently, fitness landscapes even have started to become accessible to experimental studies (Kondrashov and Kondrashov, 2015), reigniting the interest in geometric and topological features of landscapes and their relationships with properties dynamical systems operating on landscapes.

A particularly dramatic phenomenon of adaptive dynamics on landscapes is the error threshold (Eigen, 1971; McCaskill, 1984). It refers to the fact that for sufficiently small mutation rates a quasi-species, i.e., a population of sequences evolving under point-mutation and fitness proportional selection, is localized in sequence space around a peak in the fitness landscape; above a critical value of the mutation rate, however, the population becomes spread out approximately uniformly (McCaskill, 1984). The error threshold phenomenon, which is reminiscent of a phase transition, has been observed for a wide variety of fitness landscapes (Schuster, 2016), even though it does not appear to arise in all landscapes (Wiehe, 1997). As a means of studying the effects of landscape structure on dynamics in a systematic manner, tunable landscape models have been devised in which the ruggedness can be adjusted. The most widely used, and presumably best studied one is the NK model (Kauffman and Levin, 1987; Kauffman, 1989, 1993). In the realm of spin glasses,

---

W. Hordijk  
SmartAnalytiX.com E-mail: wim@WorldWideWanderings.net

S.A. Kauffman  
Institute for Systems Biology, Seattle, WA, USA; E-mail: stukauffman@gmail.com

P.F. Stadler  
Bioinformatics Group, Department of Computer Science; Interdisciplinary Center of Bioinformatics; German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig; Competence Center for Scalable Data Services and Solutions; and Leipzig Research Center for Civilization Diseases, Leipzig University, Härtelstraße 16-18, D-04107 Leipzig  
Max Planck Institute for Mathematics in the Science, Inselstraße 22, D-04103 Leipzig, Germany  
University of Vienna, Institute for Theoretical Chemistry, Währingerstraße 17, A-1090 Vienna, Austria  
Facultad de Ciencias, Universidad Nacional de Colombia, Bogotá, D.C., Colombia

the  $p$ -spin Hamiltonians play the same role (Stadler and Hapfel, 1999).

In this contribution we are concerned with the average fitness differences between a configuration  $x$  and its neighbors (Whitley *et al.*, 2008). More formally, we consider

$$(\Delta f)(x) := \sum_{y: x \sim y} (f(y) - f(x)) \quad (1)$$

The operator  $\Delta$  is (the negative of) the graph Laplacian (Mohar, 1997), which plays a key role in the algebraic theory of landscapes (Reidys and Stadler, 2002; Klemm and Stadler, 2014). The quantity that we are particularly interested in here is the expected value of  $(\Delta f)(x)$ , which measures local fitness differences, as a function of the fitness  $f(x)$  at the reference points  $x$ . Note that  $(\Delta f)(x)$  by definition measures the sum of the fitness differences between  $x$  and *all* its neighbors.

This paper is organized as follows. First, we briefly recall the NK model of fitness landscapes and some of its salient properties. We then show that the expected value of  $[(\Delta f)(x)|f(x)]$  follows a simple linear law that can be explained in terms of elementary components and the amplitude spectrum of the NK landscapes (explained in detail below). Finally, we confirm this linear relationship with numerical simulations.

## 2 The NK model and its landscapes

The *NK model* (Kauffman, 1989) serves as a general model of fitness landscapes with a tunable ruggedness introduced by epistatic interactions. It assumes a “genome” with  $N$  genes, each with two alleles (0 or 1), which is represented as a bit string  $x$  of length  $N$ . The natural neighborhood operator for bit strings (and spin glasses) is flipping a single bit or spin. The resulting NK landscape thus “lives” on a Boolean hypercube (or Hamming graphs). The fitness function  $f(x)$  of the NK model is the sum of  $N$  contributions  $f_i(x)$ , each of which depends on the  $i$ -th bit of  $x$  as well as  $K$  other bits of  $x$ , modeling epistatic interactions:

$$f(x) = \frac{1}{N} \sum_{i=1}^N f_i(x) \quad (2)$$

For  $K = 0$  there is no epistasis and  $f_i(x)$  depends only on the value of  $x_i$ .

A particular instance of the model is thus specified in terms of a table of  $N2^{K+1}$  fitness values, one for each configuration of the  $K + 1$  bits contributing to  $f_i(x)$ , together with a rule that determines which  $K$  bits influence  $f_i(x)$ . For example, the  $K$  bits can be chosen to be the nearest neighbors of  $i$ , or randomly assigned without repetitions. As we shall see later-on, the particular choice of this assignment is not relevant for our purposes.

Further details can be found in (Kauffman, 1993). The fitness values themselves are assigned from a uniform distribution. An example for  $N = 3$  and  $K = 2$  is given in the Appendix. The ruggedness of the landscape increases with  $K$ . For  $K = 0$  there are independent, additive contributions from each bit. For  $K = N - 1$ , on the other hand, there is a different random contribution  $f_i(x)$  for each value of the bit string  $x$ , hence fitness values are uncorrelated random numbers. The landscapes resulting from the NK model have been investigated in quite some detail in the literature (Weinberger, 1991; Kaul and Jacobson, 2006; Neidhart *et al.*, 2013; Buzas and Dinitz, 2014; Nowak and Krug, 2015), also in the context of an error threshold (Campos *et al.*, 2002; Ochoa, 2006).

A quite fruitful approach to analyzing landscapes starts from an algebraic representation of the underlying search space  $X$  (Stadler, 1994, 1996; Kallel *et al.*, 2001; Reidys and Stadler, 2002; Klemm and Stadler, 2014). Adjacencies on  $X$  (in our case the binary strings of length  $N$ ) are naturally defined by a “move set” (here the point mutations  $0 \leftrightarrow 1$ ). The resulting graph can be represented by its adjacency matrix  $\mathbf{A}$ , with entries  $A_{xy} = 1$  if  $x, y \in X$  are separated by a single move (point mutation), and  $A_{xy} = 0$  otherwise. The Laplacian matrix  $(-\Delta)$  has entries  $-A_{xy}$  and diagonal entries  $(-\Delta)_{xx} = \sum_{y \in X} A_{xy}$  equal to the degree of the vertex  $x$ . The negative sign is a convention connecting the formalism of the Laplacian operator in  $\mathbb{R}^n$ : discretizing space in a grid yields  $\Delta$  as the Laplacian of the grid graph.

Since the graph Laplacian  $-\Delta$  is simply an  $|X|$ -dimensional symmetric matrix, every fitness function  $f : X \rightarrow \mathbb{R}$  can be written as a linear combination  $f(x) = \sum_I a_I \varphi_I(x)$  of the eigenfunctions of  $\Delta$ . This representation is a generalization of the Fourier transformation. For the Boolean hypercube, i.e., the graph of interest here, this basis is well known: the basis functions are the (normalized) Walsh functions (Walsh, 1923)

$$\varphi_I(x) = \frac{1}{2^{N/2}} (-1)^{(I;x)} \quad (3)$$

where  $(I;x) := \sum_{j=1}^N I_j x_j$  and the “indices”  $I$  are also bit-strings of length  $N$ . The associated Laplacian eigenvalue equals  $\lambda_I = 2p(I)$ , where  $p(I)$  denotes the number of 1s in  $I$ . In particular,  $\varphi_o(x) = 1$  with  $p(o) = 0$  (i.e., eigenvalue  $\lambda_o = 0$ ) is a constant function. For all other Walsh functions we have  $\sum_{x \in X} \varphi_I(x) = 0$  and  $\sum_{x \in X} \varphi_I(x)^2 = 1$ . These properties hold true for the eigenvectors of graph Laplacians in general. For the hypercube, there are  $\binom{N}{p}$  Walsh functions of order  $p$ . This is a general phenomenon: if the graph that describes the search space is highly regular, its Laplacian matrix has many eigenvalues with high multiplicity. It therefore makes sense to collect all basis functions that belong to a given eigenvalue. These span a subspace of the fitness function. The projection of the fitness function to the eigenspace

of the Laplacian with eigenvalue  $\lambda_p$  is

$$f_p(x) = \sum_{I:\lambda_I=\lambda_p} a_I \varphi_I(x) \quad (4)$$

where the coefficients are given, as usual, by the scalar product  $a_I = \langle f, \varphi_I \rangle = \sum_{x \in X} f(x) \varphi_I(x)$ . For the hypercube the sum runs over all Walsh functions with  $|I| = p$  entries 1 in their index.

A landscape is *elementary* if it is an eigenfunction (or eigenvector) of the Laplacian up to an additive constant. More precisely, a landscape is elementary if there are constants  $\lambda$  and  $\bar{f}$  such that

$$(\Delta f)(x) = -\lambda (f(x) - \bar{f}) \quad (5)$$

where the eigenvalue  $\lambda$  is closely related to the ruggedness of the landscape (Grover, 1992; Stadler, 1994; Reidys and Stadler, 2002). The additive constant  $\bar{f}$  necessarily equals the average over the configuration space, i.e.,

$$\bar{f} := (1/|X|) \sum_{x \in X} f(x) \quad (6)$$

Thus  $f$  is elementary if (apart from the additive constant  $\bar{f}$ ) it is an eigenvector of the Laplacian, i.e., if only a single *elementary mode*  $f_p$  contributes.

Elementary landscapes have a number of appealing properties, such as a locally convex curvature near local optima, a separation of local minima and maxima by the average fitness  $\bar{f}$  (Grover, 1992), exponentially decreasing autocorrelation functions (Stadler, 1996), and a mean fitness difference at  $x$  determined by  $-\lambda(f(x) - \bar{f})/D$ , where  $D$  is vertex degree of the graph  $X$ , see e.g. (Stadler, 1996). While some landscapes, such as certain spin glasses, are elementary landscapes on the hypercube, this is not true for NK landscapes (see below).

To a certain extent, it is still possible to understand a landscape in terms of the relative contributions of the individual elementary modes, and the properties of the contributing elementary landscapes, see e.g. (Klemm and Stadler, 2014) for a recent overview. The amplitude spectrum (Stadler, 1996; Hordijk and Stadler, 1998) of the landscape describes the relative contribution of the individual elementary modes:

$$\beta_p = \langle f_p, f_p \rangle / \langle f, f \rangle = \sum_{I:p(I)=p} |a_I|^2 / \sum_{I \neq \emptyset} |a_I|^2 \quad (7)$$

for  $p = 1, \dots, N$ . Intuitively, the amplitude spectrum measures the contribution of each elementary mode to the total variance in the landscape. This is the reason why the constant contribution (with  $p(\emptyset) = 0$ ) is excluded from the definition of the amplitude spectrum. For instance, the autocorrelation function of  $f$  can be written as  $r(t) = \sum_p \beta_p r_p(t)$ , where  $r_p(t)$  is the autocorrelation function of an elementary landscape with eigenvalue  $\lambda_p$  (Stadler, 1996).

For the NK model, the amplitude spectrum can be computed in closed form independent of the particular choice of the epistatic interactions (Neidhart *et al.*, 2013):

$$\beta_p = 2^{-(K+1)} \binom{K+1}{p} \quad (8)$$

and  $\beta_p = 0$  for  $p = 0$  and  $p > K + 1$ . We note that the notation in (Neidhart *et al.*, 2013) relates by  $k = K + 1$  with  $K$  in the NK model. The latter is the number of sites with which a given site interacts in addition to its own value, while the parameter  $k$  of Neidhart *et al.* (2013) counts the total number of sites on which one of the additive contributions depends. The amplitude spectrum is independent of the choice of the interaction model.

### 3 Average fitness differences in NK landscapes

In the NK model, each additive factor contributes to the entire amplitude spectrum, i.e.,  $f = \sum_{p=0}^{K+1} f_p$ , where  $f_p$  denotes the sum term with fixed interaction order  $p$  in equ.(4). Thus

$$\Delta f = \sum_{p=0}^{K+1} \Delta f_p = \sum_{p=1}^{K+1} \lambda_p f_p \quad (9)$$

where we have used that  $\lambda_0 = 0$ . Recalling that amplitudes of the  $f_p$  are binomially distributed, this sum is dominated by values of  $p$  with the maximal values of the amplitude spectrum, i.e., the peak of the binomial distribution, or  $p \approx (K + 1)/2$ . For odd values of  $K$  this yields the approximation

$$\Delta f(x) \approx \lambda_{(K+1)/2} f = -(K + 1) f \quad (10)$$

For  $K = 0$  the landscape is elementary, hence  $\Delta f = -2f$ . For even values of  $K > 0$ , the expected amplitudes  $\beta_{K/2}$  and  $\beta_{K/2+1}$  are the same. Assuming that one dominates over the other with equal probability, or that both interaction orders contribute equally to  $f$ , we estimate  $\Delta f(x) \approx (\lambda_{K/2} + \lambda_{K/2+1})/2 \cdot f$ . We therefore arrive at the estimate

$$\Delta f \approx \begin{cases} -2f & K = 0 \\ -(K + 1)f & K \geq 1 \end{cases} \quad (11)$$

This result has not been obtained by an exact calculation but uses a rather “hand-waving” argument asserting that all but the dominating modes, i.e., the ones with the largest  $\beta_p$  contribute little to the average fitness. We therefore validate equ.(11) using numerical simulations.

To this end we picked a random starting point  $x \in X$ , generated all 1-bit neighbors, and computed  $(\Delta f)(x)/N$  and  $f(x)$ . The normalization (by the factor  $1/N$ ) was performed to compute the average different between adjacent points rather than the sum of the differences as measured by the

Laplacian operator. We then performed uphill and downhill adaptive walks to systematically sample configurations with fitness values larger and smaller than average, respectively. For different combinations of  $N$  and  $K$ , we sampled at least 200 pairs of  $f(x)$ ,  $(\Delta f)(x)$  values. Data were generated for  $N = 10, 20, 30, 40, 50$  and  $K = 0, 5, 9, 19, 29, 39, 49$  with  $K < N$ .

Figure 1 shows a scatter plot of  $(\Delta f)(x)/N$  against  $f(x)$ . Both panels show that for each value of  $K$  the average  $\Delta f|f^0$  of  $\Delta f$  over strings with approximate fitness  $f^0$  exhibits the expected linear dependence on  $f^0$ . Also following our approximation, the slope  $s$  for fixed  $N$  increases with  $K$  in Figure 1(a). Both panels show that the average fitness difference  $\Delta f|f^0/N$  vanishes for  $f^0 = 1/2$ , the expected fitness value of a randomly chosen bit string in an NK model.

The empirically determined slopes are summarized in Fig. 2(a) for different values of  $N$  and  $K$ . Normalizing Equ.(11) by the factor  $1/N$  yields the following predictions for the slopes:  $s = 2/N$  for  $K = 0$  and  $s = (K + 1)/N$  for all other values of  $K$ . Fig. 2(b) shows that the theoretical and empirical values are in excellent agreement. In particular we see that the “mixed order” NK landscapes with  $K > 0$  behave differently from the elementary (additive) landscape for  $K = 0$ . The slope  $s$  is completely determined by  $N$  and  $K$ . Since the autocorrelation function and the amplitude spectrum is influenced neither by the parity of  $K$  nor the choice of the adjacency relation among the sites (Neidhart *et al.*, 2013),  $s$  is also unaffected by these parameters. Some empirical results to confirm this are provided in the Appendix. In contrast, the length of adaptive walks and the distribution of local optima depends the interaction scheme (Buzas and Dinitz, 2014; Nowak and Krug, 2015). As in the case of spin-glass models, the dominating influence is exerted by the rank, that is, the number of (linearly) independent, non-zero contributions  $a_l$  in equ.(4) (de Oliveira *et al.*, 1999).

## 4 Discussion

The average fitness differences on NK landscapes depend in a simple manner on the the number of bits/spins  $N$ , the strength of epistasis  $K$ , and the fitness of the reference point. These dependencies are easily understood in terms of the elementary modes, with the dominating mode, which corresponds to the interaction of  $(K + 1)/2$  sites, determining the behavior, except for the special case  $K = 0$ . The parameter  $K$  in the NK model serves as a measure of ruggedness. To see this, recall that the correlation length can be computed from the amplitude spectrum (Stadler, 1996). For the NK model

we obtain

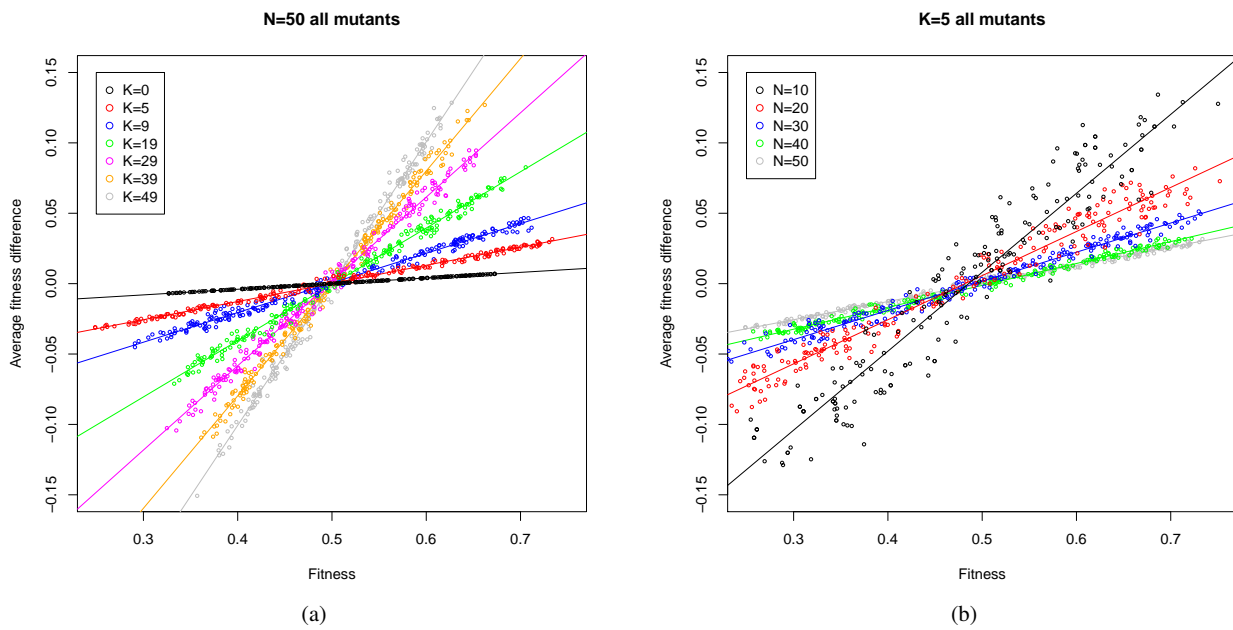
$$\begin{aligned} \ell &= N \sum_{p>0} \beta_p / \lambda_p = \frac{N}{2^{K+2}} \sum_{p=0}^{K+1} \frac{1}{p} \binom{K+1}{p} \\ &\approx \frac{N}{2(K+1)} \end{aligned} \quad (12)$$

for large  $K$ . The slope  $s$  of the average fitness differences thus coincides with the inverse correlation length. As expected intuitively, therefore, we have larger local fitness variations on more rugged, i.e., less correlated, landscapes.

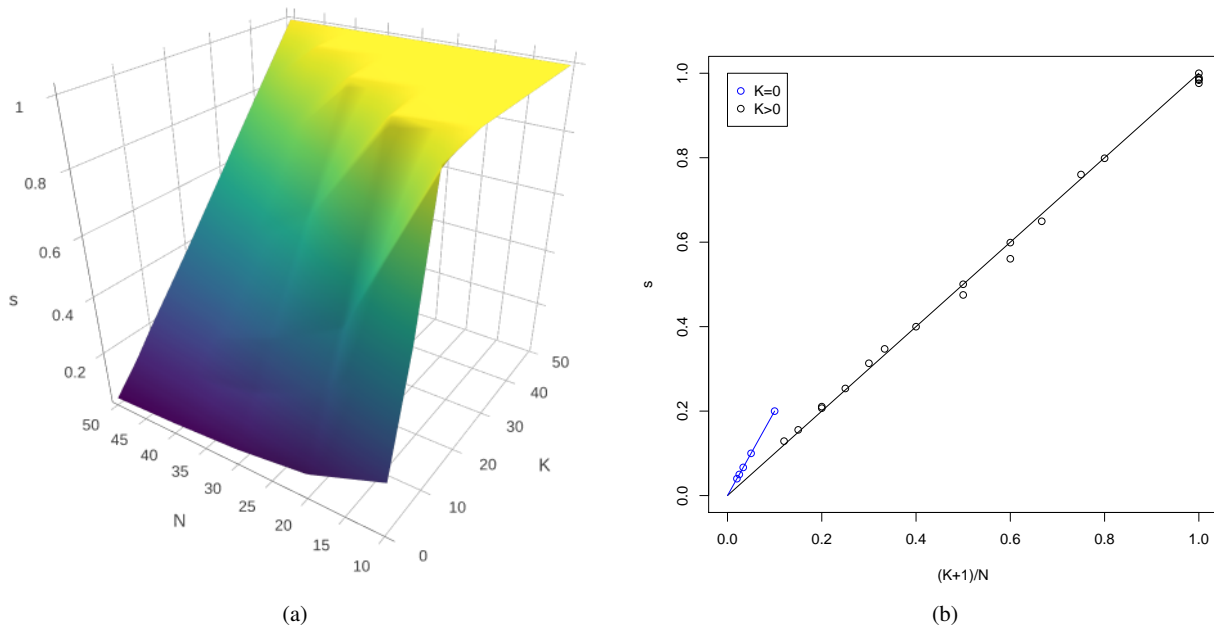
These observations have direct consequences for the possible existence of an error threshold on NK landscapes. Schuster observed that “...a smaller difference in fitness other factors being unchanged causes the transition to occur at a smaller [value of the mutation rate].” (Schuster, 2016, p.100). More precisely, the error threshold, that is, the maximally tolerable mutation rate  $p_{\max}$  per digit and replication satisfies  $Np_{\max} \approx \ln \sigma$ , where the *superiority*  $\sigma$  is a population-averaged fitness difference (Eigen *et al.*, 1989). The population-independent local average fitness difference  $(\Delta f)/N$  is only an approximation for  $\sigma$ , which nevertheless reflects the general trends very well (Bonhoeffer and Stadler, 1993; Semenov *et al.*, 2014).

Consequently, on smoother landscapes the error threshold will already occur at lower mutation rates  $\mu$ , while on more rugged landscapes it will only occur at higher  $\mu$ . So, even though it is more difficult to find a high-fitness peak on a more rugged landscape, once one is found it is more likely that the population can stay on it compared to smoother landscapes, where high-fitness peaks are easier to find but harder to maintain. Our results provide precise quantitative measures for such comparisons.

An important caveat is this that this statement is true of landscapes that are sufficiently “isotropic” (García-Pelayo and Stadler, 1997; Stadler and Happel, 1999) in the sense that a large number of individual degrees of freedom (coefficients  $a_l$  in equ.(4)) contribute with comparable magnitudes to each mode  $\beta_p$  of the amplitude spectrum, and that the amplitude spectrum itself is dominated by a small number of modes with similar eigenvalues, i.e., adjacent values of the interaction order  $p$ . The first property ensures that there is not a lot of neutrality in the landscapes (Reidys and Stadler, 2001). The second property ensures that the landscape behaves similar to an elementary landscapes (Stadler, 1996), in which case the linear relationship between  $(\Delta f)/N$  and  $f$  is exact. Landscape that drastically deviate from these “benign” properties may behave very differently, and in extreme cases, do not show a sharp error threshold at all (Wiehe, 1997; Bull *et al.*, 2005).



**Fig. 1** Scatter plot of  $(\Delta f)(x)/N$  against  $f$ . (a) Data for  $N = 50$  and different values of  $K$ . (b) Data for  $K = 5$  and various values of  $N$ . Straight lines represent linear regression fits.



**Fig. 2** Slopes  $s$  as a function of  $N$  and  $K$ . (a) 3D plot. For  $K \geq N$  the slope  $s$  is not defined, but for visualization purposes these slopes are set to  $s = 1$ . Also note that the “folds” in this figure are an artifact of using only a limited number of  $K$  values per  $N$  value. In other words, the “ridge” should be straight rather than jagged. An interactive version of this plot, which can be zoomed and rotated, can be accessed at <http://worldwidewanderings.net/General/NK/html/index.html>. (b) Empirically determined slopes as a function of  $(K+1)/N$ , drawn separately for  $K = 0$  (blue) and  $K \geq 1$  (black), respectively. Within the stochastic uncertainty, measured (dots) and predicted (lines) slopes are in agreement.

## Acknowledgments

WH thanks the Institute for Advanced Study, Amsterdam, for financial support through a fellowship. Discussions with Bärbel M. R. Stadler are gratefully acknowledged.

## References

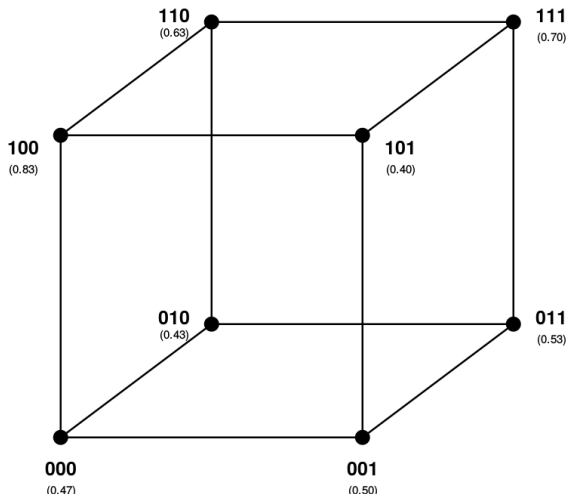
Bonhoeffer S, Stadler PF, 1993. Error threshold on complex fitness landscapes. *J Theor Biol* 164:359–372.

- Bull JJ, Ancel Meyers L, Lachmann M, 2005. Quasispecies made simple. *PLoS Comput Biol* 1:e61.
- Buzas J, Dinitz J, 2014. An analysis of NK landscapes: Interaction structure, statistical properties, and expected number of local optima. *IEEE Trans Evol Comput* 18:807–818.
- Campos PRA, Adami C, Wilke CO, 2002. Optimal adaptive performance and delocalization in NK fitness landscapes. *Physica A* 304:495–506.
- de Oliveira VM, Fontanari JF, Stadler PF, 1999. Metastable states in high order short-range spin glasses. *J Phys A Math Gen* 32:8793–8802.
- Eigen M, 1971. Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften* 10:465–523.
- Eigen M, McCaskill J, Schuster P, 1989. The molecular quasispecies. *Adv Chem Phys* 75:149–263.
- García-Pelayo R, Stadler PF, 1997. Correlation length, isotropy, and meta-stable states. *Physica D* 107:240–254.
- Grover LK, 1992. Local search and the local structure of NP-complete problems. *Oper Res Lett* 12:235–243.
- Hordijk W, Stadler PF, 1998. Amplitude spectra of fitness landscapes. *Journal of Complex Systems* 1:39–66.
- Kallel L, Naudts B, Reeves CR, 2001. Properties of fitness functions and search landscapes. In: Kallel L, Naudts B, Rogers A, editors, *Theoretical Aspects of Evolutionary Computing*, (pp. 175–206). Heidelberg: Springer.
- Kauffman SA, 1989. Adaptation on rugged fitness landscapes. In: Stein D, editor, *Lectures in the Sciences of Complexity*, (pp. 527–618). Addison-Wesley.
- Kauffman SA, 1993. *Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press.
- Kauffman SA, Levin S, 1987. Towards a general theory of adaptive walks on rugged landscapes. *J Theor Biol* 128:11–45.
- Kaul H, Jacobson SH, 2006. Global optima results for the kauffman NK model. *Mathematical Programming* 106:319–338.
- Klemm K, Stadler PF, 2014. Rugged and elementary landscapes. In: Borenstein Y, Moraglio A, editors, *Theory and Principled Methods for Designing Metaheuristics*, Natural Computing Series, (pp. 41–61). Berlin: Springer.
- Kondrashov DA, Kondrashov FA, 2015. Topological features of rugged fitness landscapes in sequence space. *Trends in Genetics* 31:24–33.
- McCaskill JS, 1984. A localisation threshold for macromolecular quasispecies from continuously distributed replication rates. *J Chem Phys* 80:5194–5202.
- Mohar B, 1997. Some applications of Laplace eigenvalues of graphs. In: Hahn G, Sabidussi G, editors, *Graph Symmetry: Algebraic Methods and Applications*, vol. 497 of *NATO ASI Series C*, (pp. 227–275). Dordrecht: Kluwer.
- Neidhart J, Szendro IG, Krug J, 2013. Exact results for amplitude spectra of fitness landscapes. *Journal of Theoretical Biology* 332:218–227.
- Nowak S, Krug J, 2015. Analysis of adaptive walks on NK fitness landscapes with different interaction schemes. *J Stat Mech* (p. P06014).
- Ochoa G, 2006. Error thresholds in genetic algorithms. *Evol Comp* 14:157–182.
- Reidys CM, Stadler PF, 2001. Neutrality in fitness landscapes. *Appl Math Comput* 117:321–350.
- Reidys CM, Stadler PF, 2002. Combinatorial landscapes. *SIAM Review* 44:3–54.
- Richter H, Engelbrecht A, editors, 2014. *Recent advances in the theory and application of fitness landscapes*, vol. 6 of *Emergence, Complexity, and Computation*. Berlin: Springer-Verlag.
- Schuster P, 2016. Quasispecies on fitness landscapes. In: Domingo E, Schuster P, editors, *Quasispecies: From Theory to Experimental Systems*, (pp. 61–120). Heidelberg: Springer.
- Semenov YS, Bratus AS, Novozhilov AS, 2014. On the behavior of the leading eigenvalue of Eigen’s evolutionary matrices. *Math Biosci* 258:134–147.
- Stadler BMR, Stadler PF, 2002. Generalized topological spaces in evolutionary theory and combinatorial chemistry. *J Chem Inf Comput Sci* 42:577–585.
- Stadler PF, 1994. Linear operators on correlated landscapes. *J Phys I France* 4:681–696.
- Stadler PF, 1996. Landscapes and their correlation functions. *Journal of Mathematical Chemistry* 20:1–45.
- Stadler PF, Happel R, 1999. Random field models for fitness landscapes. *J Math Biol* 38:435–478.
- Walsh JL, 1923. A closed set of normal orthogonal functions. *Amer J Math* 45:5–24.
- Weinberger E, 1991. Local properties of Kauffman’s N-k model: A tunably rugged energy landscape. *Physical Review A* 44:6399–6413.
- Whitley LD, Sutton AM, Howe AE, 2008. Understanding elementary landscapes. In: Ryan C, Keijzer M, editors, *Genetic and Evolutionary Computation Conference, GECCO 2008*, (pp. 585–592). ACM.
- Wiehe T, 1997. Model dependency of error thresholds: the role of fitness functions and contrasts between the finite and infinite sites models. *Genetics Research* 69:127–136.
- Wright S, 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. In: Jones DF, editor, *Proceedings of the Sixth International Congress on Genetics*, vol. 1, (pp. 356–366). New York: Brooklyn Botanic Gardens.
- Wright S, 1967. “Surfaces” of selective value. *Proc Nat Acad Sci USA* 58:165–172.

## A An example of an NK landscape

We include here a brief example of an NK landscape, taken from (Kauffman, 1993). The fitness contributions  $f_i(b)$  for the individual positions are tabulated for each value of the bit  $b_i$  itself and  $K = 2$  additional relevant bits. The final fitness value for each bit string is the average of the position-wise contributions. Together with the adjacency relation of the Boolean hypercube, this defines the landscape.

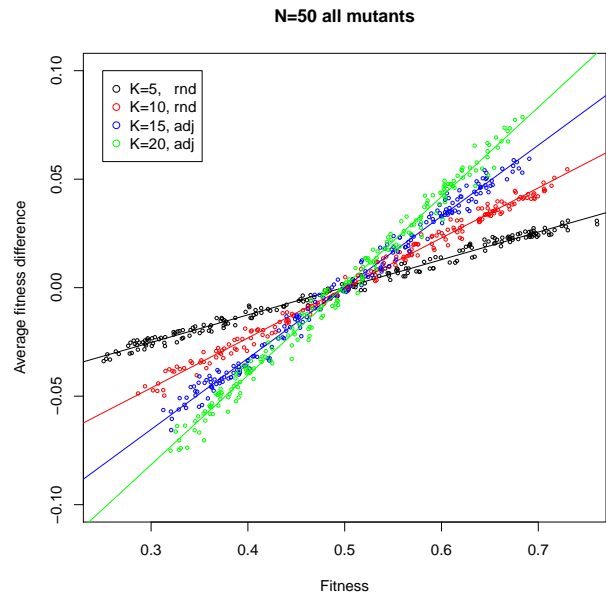
value of bit			fitness contribution			total fitness
$b_1$	$b_2$	$b_3$	$f_1$	$f_2$	$f_3$	$f = \frac{1}{N} \sum_{i=1}^N f_i$
0	0	0	0.6	0.3	0.5	0.47
0	0	1	0.1	0.5	0.9	0.50
0	1	0	0.4	0.8	0.1	0.43
0	1	1	0.3	0.5	0.8	0.53
1	0	0	0.9	0.9	0.7	0.83
1	0	1	0.7	0.2	0.3	0.40
1	1	0	0.6	0.7	0.6	0.63
1	1	1	0.7	0.9	0.5	0.70



**Fig. 3** A simple example of an instance of the NK model for  $N = 3$  and  $K = 2$ . Top: The fitness contributions for the three bits for each of the  $2^{K+1} = 8$  possible neighborhood configurations are assigned at random. The fitness of the entire string is the average of the individual fitness contributions. Bottom: The boolean hypercube representing the fitness landscape defined in the table above.

## B Independence of Parity and Neighborhood Structure

Figure 4 shows a scatter plot of  $(\Delta f)(x)/N$  against  $f$  for landscapes with different parity and different neighborhood structures (random and adjacent). We have chosen the different values of  $K$  such that the data sets are distinguishable. The predicted slopes for  $N = 50$  and  $K = 5, 10, 15,$  and  $20$  are  $s = 0.12, 0.22, 0.32,$  and  $0.42$ , while the empirical values from the data displayed here are  $\hat{s} = 0.127, 0.230, 0.328,$  and  $0.412$ , respectively. The empirical and theoretical values are in excellent agreement.



**Fig. 4** Scatter plot of  $(\Delta f)(x)/N$  against  $f$ , illustrating that the slopes do not depend on parity, i.e., even vs. odd values of  $K$ , or the choice of adjacency, i.e., random ('rnd') vs. adjacent ('adj').