# ESTIMATING SOME FEATURES OF NK FITNESS LANDSCAPES

STEVEN N. EVANS AND DAVID STEINSALTZ

ABSTRACT. Kauffman and Levin (1987) introduced a class of models for the evolution of hereditary systems which they called "$NK$ fitness landscapes". Inspired by spinglasses, these models have the attractive feature of being tunable, with regard to both overall size (through the parameter $N$) and connectivity (through $K$). There are $N$ genes, each of which exists in two possible alleles (leading to a system indexed by $\{0,1\}^N$); the fitness score of an allele at a given site is determined by the alleles of $K$ neighboring sites. Otherwise the fitnesses are as simple as possible, namely i.i.d., and the fitnesses of different sites are simply averaged.

Much attention has been focused on these fitness landscapes as paradigms for investigating the interaction between size and complexity in making evolution possible. In particular, the effect of the interaction parameter $K$ on the height of the global maximum and the heights of local maxima has attracted considerable interest, as well as the behavior of a "hill-climbing" walk from a random starting point. Nearly all of this work has relied on simulations, not on rigorous mathematics.

In this paper, some asymptotic features of $NK$ fitness landscapes are reduced to questions about eigenvalues and Lyapunov exponents. When $K$ is fixed, the expected number of local maxima grows exponentially with $N$ at a rate depending on the top eigenvalue of a kernel derived from the distribution of the fitnesses, and the average height of a local maximum converges to a value determined by the corresponding eigenfunction.

The global maximum converges in probability as $N \to \infty$ to a constant given by the top Lyapunov exponent for a system of i.i.d. max–plus random matrices, and this constant is non-decreasing with $K$. Various such quantities are computed for certain special cases when $K$ is small, and these calculations can, in principle, be extended to larger $K$.

**Short title: $NK$ Fitness**

## 1. INTRODUCTION

1.1. **Background.** Early in the twentieth century Sewall Wright (see, for example, [Wri32]) proposed what has become one of the dominant metaphors in the analysis of biological evolution: the fitness landscape. Biological evolution is modelled as

*Date*: February 8, 2002.

1991 *Mathematics Subject Classification.* 92D15, 60G60, 60G70, 47A75.

*Key words and phrases.* genetics, evolution, eigenvalue, Perron–Frobenius, Lyapunov exponent, max–plus algebra, random field, spinglass, extreme value.

Evans supported in part by NSF grant DMS-0071468.

gradual motion through an abstract space, which represents the possible genomes and other heritables. To every point in this space is assigned a number, the "fitness function", that summarizes the relative success of an organism with this particular endowment in the struggle for existence. The graph of this fitness function, with its narrow peaks of high functioning and broad valleys of dissolute DNA, is conceived as being a "fitness landscape". Natural selection appears as a random walk of a species along this landscape, with a bias toward upward steps.

The fitness landscape, with its associated diagrams, can be a misleading image. The landscape is viewed as stable, though of course coevolution of species and their environments is universal. The space of possibilities itself is not fixed, as the machinery of heredity evolves. And the fitness of a particular genome depends very much on the number and type of its conspecifics, a complication concealed by the naive picture of a solitary point scaling its Sierras of success.

Even closing an eye to all of these defects, we see something basically deficient about the intuitions that the fitness-landscape story conjures up. One tends to imagine a smooth, undulating, two-dimensional terrain: here a bit fitter, there a bit less fit. But even the crudest models of genetic space are nothing at all like Euclidean space. They are high-dimensional discrete spaces, such as the "DNA space" $\{A, C, G, T\}^L$, or the "Mendelian space" $\{0, 1, \ldots, r-1\}^N$, representing the possible genomes when there are $N$ genetic loci, each with $r$ possible alleles. There will be only slivers of high fitness looming up above the vast genomic tohubohu. An evolving organism is not roaming through the whole space, but creeping along these fitness spines. Neighborhoods are defined by small Hamming distance, meaning differences in a small number of alleles, but synergies may give these small changes enormous effect. From any given point a small step in most directions is a calamity: the fitness landscape is, in the vernacular, "rugged". How fit will the organism get, seeking its highest level in such a landscape?

The answer may or may not depend on the precise details of the model, but a rigorous answer does require at the least a model that is rigorously defined. One class of models that has received a significant amount of attention in recent years

is Kauffman and Levin's $NK$ *fitness landscape* [KL87, Kau93]. Based on spinglass models in statistical physics, this is a stochastic process indexed by the Mendelian space with two possible alleles at each of $N$ loci. The quantity $K$ is an interaction parameter that tunes the ruggedness. Each locus is assumed to rely for its fitness on $K$ other loci. Beyond that, fitness is random; that is, each of the $2^{K+1}$ possible assignments of alleles to the gene and its entourage gets an independent random value. Since these sets overlap, there is no easy way to find the optimal choice for all $N$ alleles: a choice that improves one fitness will likely detract from another.

While no one would mistake this abstract system for a realistic model of genetic evolution, it has the virtues of a good foundational model: It is easy to describe, yet contains a wealth of structure that is neither obvious nor superficially accessible. Before we can analyze a more realistic model, it would seem that we must first come to grips with models such as this one. At the same time, we may hope that some general features of this model will carry over to something like the real world.

1.2. **The Model.** We begin with the genetic space $S = \{0,1\}^N$. A fitness function is a map from $S$ to the real numbers. Gene interactions are confined to a range of $K$: a gene interacts with its $K$ successors, and the fitness contribution of allele number $i$ and its $K$ successors is given by a function $F_i : \{0,1\}^{K+1} \to \mathbb{R}$. Successors are defined cyclically, that is, modulo $N$. The contributions of different genes are assumed additive, so that the total fitness function is

$$G_{N,K}(x_0, \ldots, x_{N-1}) = \frac{1}{N} \sum_{i=0}^{N-1} F_i(x_i, \ldots, x_{i+K}).$$

This differs from Kauffman's model in two small ways. Kauffman considered two versions of the interaction, one in which a locus's $K$ neighbors were determined independently at random, and the other where the interaction was confined to a range of $K/2$ on either side. For $K$ even our process is equivalent to the latter; our choice to put the neighbors all on one side was purely a matter of notational convenience. Kauffman also performed his simulations where the fitness distribution was uniform on the interval $[0,1]$, while conjecturing that much would remain unchanged if this were replaced by another distribution. While we have some things to say about

general distributions for the $F_i$, we have had the most success in obtaining explicit quantitative information when the $F_i$ are exponentially distributed.

Kauffman was primarily concerned to ask, if an imaginary organism starts in a random place in this vast genomic space, and then walks upward to the nearest local maximum, how far will it get? How far will the fitness be above average? Clearly in the case $K = 0$, where fitnesses of different loci are completely uncorrelated, there is only one local maximum, so the process eventually will reach the global maximum. At the other extreme, when $K = N - 1$, the global maximum will be larger, but it will never be reached, since local maxima will be ubiquitous. In Kauffman's computer simulations with small values of $N$ (generally up to $N = 96$) the local maxima actually attained were largest for small values of $K$, around 4.

Other work, such as [Wil98] and [SBTG99], attempts to generalize the model, admitting in the one case time-dependent fitness landscapes, in the other non-i.i.d. fitnesses and fitnesses whose standard-deviations decrease with $N$ (to illustrate how the "complexity catastrophe", the tendency of the local maxima to collapse to mere average behavior as $K$ increases, may be mitigated.) These results rely primarily on simulations, hence on small values of $N$, to make their points. (One paper [SJ94] does present simulations for large values of $N$, facilitated by making the $F_i$'s Bernoulli variables.) The only exception we have found is the paper [Wei91], which purports to derive asymptotic formulae for the number of local maxima, when $N$ and $K$ are large and the fitnesses are normally distributed. However, it appears that at a crucial step an error term is discarded which seems to dominate the favored approximation.

Independent work related to ours appears in [DL01]. We will describe the connection between that paper and our work in the course of outlining our results in the next section.

1.3. **Outline of results.** We begin in Section 2 by presenting some general results about the global maximum. Lemma 1 connects the maximum to the max–plus "product" of certain random matrices. Unrelated to this, we then show in Proposition 2 that the maximum is stochastically non-decreasing in $K$, in the special case

when the fitnesses are Gaussian. In Section 3 we show that, for $K$ fixed, the global maximum converges in probability to a constant as $N \to \infty$, and that this constant is the solution of a certain variational problem. This *a priori* infinite dimensional variational problem turns out to be finite dimensional when the fitnesses have exponential distributions. Explicit numerical computations are therefore "just" a matter of finite dimensional linear algebra, although these calculations quickly grow infeasible as $K$ increases. We carry through the particularly tractable example of $K = 1$. We also establish for general fitness distributions that the asymptotic value of the global maximum is non-decreasing in $K$.

In Section 4 we show, under the assumption that the support of the fitness distribution is bounded below, that the expected number of local maxima increases geometrically in $N$, by a power that is computable, in theory, as the spectral radius of a certain operator derived from the fitness distribution. Section 5 applies the same principles to the limit of the expected height of a typical local maximum. In both sections the computations are actually carried out for the special cases $K = 1$ and $K = 2$, where the fitnesses are exponentially distributed. For exponentially distributed fitnesses (and, more generally, fitnesses distributed according to a gamma distribution with integer shape parameter), the *a priori* infinite dimensional problem of determining the spectral radius reduces to a finite dimensional one. However, the dimension grows rapidly with $K$.

The height of a typical local maximum is investigated in [DL01] using tools that differ from those used here (primarily the theory of $R$-recurrent Markov chains). In particular, explicit calculations are carried out in [DL01] for the case where $K = 1$ and the fitnesses are the negatives of exponentially distributed random variables. Our result for $K = 1$ and exponentially distributed fitnesses with mean 1 is that the expected height of a typical local maximum converges to 1.61651 as $N \to \infty$, whereas a result in [DL01] implies that the expected height of typical local *minimum* for the same model converges to 0.480971. A central limit theorem for the height of a typical local maximum and a large deviation result for the global maximum are also given in [DL01].

Section 6 offers an alternative representation of the probability for a point to be a local maximum, in the exponential case, with general $K$. Not only does this representation provide a way of estimating the expected number of local maxima for finite $N$ via a Monte–Carlo approach that is considerably simpler than simulating the model itself and then determining which points are local maxima, but it also provides an interesting coupling between models with different values of $K$. This coupling may be useful for further analytic investigations of the dependence on $K$ of the expected number of local maxima.

Section 7 states a version of the Perron–Frobenius Theorem for infinite–dimensional operators, which we use in Sections 4 and 5. We did not find this result in the form we needed in the literature, but it can be proved by a fairly straightforward adaptation of the classical result for matrices.

## 2. The global maximum

We define $2^K \times 2^K$ $\mathbb{R} \cup \{-\infty\}$-valued matrices indexed by $\{0,1\}^K$:

$$A_i(x_0, \dots, x_{K-1}; y_0, \dots, y_{K-1}) := \begin{cases} F_i(x_0, \dots, x_{K-1}, y_{K-1}) & \text{if } (x_1, \dots, x_{K-1}) = (y_0, \dots, y_{K-2}), \\ -\infty & \text{otherwise.} \end{cases}$$

For the moment, the fitnesses $F_i$ are just arbitrary functions on $\{0,1\}^{K+1}$. We view these as matrices in the $(\vee, +)$ algebra (where $a \vee b$ denotes the maximum of $a$ and $b$), so that the product $A_0 A_1$ is given by

$$A_0 A_1(\vec{x}, \vec{y}) := \bigvee_{\vec{z} \in \{0,1\}^K} \Big\{ A_0(\vec{x}, \vec{z}) + A_1(\vec{z}, \vec{y}) \Big\}.$$

We denote the maximum fitness by

$$G^*_{N,K} := \bigvee_{(x_0, \dots, x_{N-1}) \in S} G_{N,K}(x_0, \dots, x_{N-1}).$$

**Lemma 1.**

$$G^*_{N,K} = \frac{1}{N} \bigvee_{\vec{x} \in \{0,1\}^K} A_0 A_1 \cdots A_{N-1}(\vec{x}, \vec{x}).$$

*Proof.* Let

$$M_n = A_0 A_1 \cdots A_n$$

for $0 \le n \le N-1$. We prove by induction that

$$M_n(\vec{x},\vec{y}) = \bigvee\Big\{\sum_{i=0}^{n} F_i(z_i,\dots,z_{i+K}) \ : z_i = x_i \text{ for } 0 \le i \le K-1 \text{ and } z_{n+1+i} = y_i \text{ for } 0 \le i \le K-1\Big\}. \tag{1}$$

For $n = 0$ the statement is merely the definition of $A_0$. Suppose it is true for $n-1$. Let $G_n^*(\vec{x},\vec{y})$ be the right-hand side of (1). Then

$$\begin{aligned} M_n(\vec{x},\vec{y}) &= \big(G_{n-1}^* \cdot A_n\big)(\vec{x},\vec{y}) \\ &= \bigvee\Big\{G_{n-1}^*(\vec{x},(0,y_0,\dots,y_{K-2})) + F_n(0,y_0,\dots,y_{K-1}) \,, \\ &\qquad G_{n-1}^*(\vec{x},(1,y_0,\dots,y_{K-2})) + F_n(1,y_0,\dots,y_{K-1})\Big\} \\ &= G_n^*(\vec{x},\vec{y}). \end{aligned}$$

□

For the rest of the paper, take the $F_i(\vec{x})$'s to be given by $N \cdot 2^{K+1}$ i.i.d. random variables, with distribution function $\mathcal{F}$. (We will use $\mathcal{F}_*$ to denote the distribution function of the sum of $K+1$ i.i.d. copies of a random variable with distribution $\mathcal{F}$.)

**Proposition 2.** *If the fitnesses are normal variables, the maximum is stochastically non-decreasing in $K$. That is, for any fixed $N$ and any real number $z$,*

$$\mathbb{P}\big\{G_{N,K+1}^* \ge z\big\} \ge \mathbb{P}\big\{G_{N,K}^* \ge z\big\}. \tag{2}$$

*Proof.* By the obvious linearity, we may as well assume that the fitnesses are standard normal. The average fitness $G_{N,K}(\vec{x})$ is a Gaussian process, indexed by $\{0,1\}^N$. For any $\vec{x},\vec{y} \in S$, the covariance of $G_{N,K}(\vec{x})$ and $G_{N,K}(\vec{y})$ is

$$\begin{aligned} &\frac{1}{N^2}\sum_{i=0}^{N-1} \mathbb{E}\big[F_i(x_i,\dots,x_{i+K})F_i(y_i,\dots,y_{i+K})\big] \\ &\qquad = \frac{1}{N^2}\#\big\{i \ : (x_i,\dots,x_{i+K}) = (y_i,\dots,y_{i+K})\big\}. \end{aligned}$$

If $i$ is an index such that $(x_i, \ldots, x_{i+K+1}) = (y_i, \ldots, y_{i+K+1})$, then $(x_i, \ldots, x_{i+K}) = (y_i, \ldots, y_{i+K})$. The covariance of $G_{N,K}(\vec{x})$ and $G_{N,K}(\vec{y})$ is is thus nonincreasing in $K$. Hence, the Gaussian processes $G_{N,K}$ and $G_{N,K+1}$ have the same expectations and variances at each point, but the covariances of $G_{N,K+1}$ are smaller for any pair of points. The inequality (2) is then an immediate consequence of Slepian's Lemma [Kah85, Section 15.2]. □

We will show in Section 3 below that, for each fixed $K$, $G^*_{N,K}$ converges in probability as $N \to \infty$ to a constant $\lambda_K$. The following general asymptotic result complements the special finite $N$ result of Proposition 2.

**Proposition 3.** *For an arbitrary fitness distribution $\mathcal{F}$, the asymptotic global maximum $\lambda_K$ is non-decreasing in $K$.*

*Proof.* Fix $K$. To emphasize the rôle of $K$, let $A_0^K, A_1^K, \ldots$ be the $2^K \times 2^K$ matrices defined above for the $NK$ model, and let $A_0^{K+1}, A_1^{K+1}, \ldots$ be the analogous $2^{K+1} \times 2^{K+1}$ matrices for the $N(K+1)$ model. From the proof of Lemma 1 and Section 3 below we see for any fixed $L > K$ that the matrices $A_0^K, A_1^K, \ldots, A_{N-L}^K$ are i.i.d. and that

$$\lambda_K = \lim_{N\to\infty} \frac{1}{N} \bigvee \{A_0^K A_1^K \cdots A_{N-L}^K(\vec{x}, \vec{y}) : \vec{x}, \vec{y} \in \{0,1\}^K\}.$$

A similar remark holds with $K$ replaced by $K+1$. In order to show that $\lambda_{K+1} \geq \lambda_K$ it certainly suffices to show for a fixed $L > K+1$ and all $N \geq L$ that

$$\bigvee \{A_0^{K+1} A_1^{K+1} \cdots A_{N-L}^{K+1}(\vec{x}, \vec{y}) : \vec{x}, \vec{y} \in \{0,1\}^{K+1}\}$$

stochastically dominates

$$\bigvee \{A_0^K A_1^K \cdots A_{N-L}^K(\vec{x}, \vec{y}) : \vec{x}, \vec{y} \in \{0,1\}^K\}.$$

Using the fitnesses $F_i$ for the $NK$ model define new $2^{K+1} \times 2^{K+1}$ $\mathbb{R} \cup \{-\infty\}$-valued matrices indexed by $\{0,1\}^{K+1}$ by

$$\tilde{A}_i^{K+1}(x_0, \ldots, x_K; y_0, \ldots, y_K)$$

$$:= \begin{cases} F_i(x_0, \ldots, x_{K-1}, y_{K-1}) & \text{if } (x_1, \ldots, x_{K-1}) = (y_0, \ldots, y_{K-2}), \\ -\infty & \text{otherwise.} \end{cases}$$

It is not hard to see that

$$\bigvee \{A_0^K A_1^K \cdots A_{N-L}^K(\vec{x}, \vec{y}) : \vec{x}, \vec{y} \in \{0,1\}^K\}$$
$$= \bigvee \{\tilde{A}_0^{K+1} \tilde{A}_1^{K+1} \cdots \tilde{A}_{N-L}^{K+1}(\vec{x}, \vec{y}) : \vec{x}, \vec{y} \in \{0,1\}^{K+1}\}.$$

It thus suffices to show that

$$\bigvee \{A_0^{K+1} A_1^{K+1} \cdots A_{N-L}^{K+1}(\vec{x}, \vec{y}) : \vec{x}, \vec{y} \in \{0,1\}^{K+1}\}$$

stochastically dominates

$$\bigvee \{\tilde{A}_0^{K+1} \tilde{A}_1^{K+1} \cdots \tilde{A}_{N-L}^{K+1}(\vec{x}, \vec{y}) : \vec{x}, \vec{y} \in \{0,1\}^{K+1}\},$$

which in turn will follow if we can establish for $0 \leq k \leq N - L - 1$ that

$$\bigvee \{A_0^{K+1} A_1^{K+1} \cdots A_k^{K+1} \tilde{A}_{k+1}^{K+1} \cdots \tilde{A}_{N-L}^{K+1}(\vec{x}, \vec{y}) : \vec{x}, \vec{y} \in \{0,1\}^{K+1}\}$$

stochastically dominates

$$\bigvee \{A_0^{K+1} A_1^{K+1} \cdots A_{k-1}^{K+1} \tilde{A}_k^{K+1} \cdots \tilde{A}_{N-L}^{K+1}(\vec{x}, \vec{y}) : \vec{x}, \vec{y} \in \{0,1\}^{K+1}\}.$$

This, however, follows from the independence of the matrices in the products and Lemma 4 below with $m = 2^{K+1}$, $\{1, \ldots, m\}$ identified with $\{0,1\}^{K+1}$, and the partition $\Pi$ of $\{0,1\}^{K+1} \times \{0,1\}^{K+1}$ such that two pairs of indices $(x_0, \ldots, x_K; y_0, \ldots, y_K)$ and $(x'_0, \ldots, x'_K; y'_0, \ldots, y'_K)$ are in the same block of $\Pi$ if and only if

$$(x_0, \ldots, x_{K-1}; y_0, \ldots, y_{K-1}) = (x'_0, \ldots, x'_{K-1}; y'_0, \ldots, y'_{K-1}).$$

□

**Lemma 4.** *Let $B$ be a random $m \times m$ $\mathbb{R} \cup \{-\infty\}$-valued matrix with independent entries. Suppose that $\Pi$ is a partition of $\{1, \dots, m\} \times \{1, \dots, m\}$ with the property that if $(i,j)$ and $(k,\ell)$ belong to the same block of $\Pi$, then $B_{ij}$ and $B_{k\ell}$ have the same distribution. Let $\tilde{B}$ be another random $m \times m$ $\mathbb{R} \cup \{-\infty\}$-valued matrix with the properties:*

- *For each pair $(i,j)$, $B_{ij}$ and $\tilde{B}_{ij}$ have the same distribution.*
- *If $(i,j)$ and $(k,\ell)$ belong to the same block of $\Pi$, then $\tilde{B}_{ij} = \tilde{B}_{k\ell}$.*
- *If $, _1, \dots, , _n$ are the blocks of $\Pi$, then the collections of random variables $\{\tilde{B}_{ij} : (i,j) \in , _h\}$, $1 \le h \le n$, are independent.*

*Then for any two constant $m \times m$ $\mathbb{R} \cup \{-\infty\}$-valued matrices $A$ and $C$ the random variable $\bigvee\{(ABC)_{ij} : 1 \le i,j \le m\}$ stochastically dominates the random variable $\bigvee\{(A\tilde{B}C)_{ij} : 1 \le i,j \le m\}$.*

*Proof.* By truncation and taking limits, we may suppose that all matrices are real–valued.

We need to show that

$$\mathbb{P}\left\{\bigvee\{(ABC)_{ij} : 1 \le i,j \le m\} \le x\right\} \le \mathbb{P}\left\{\bigvee\{(A\tilde{B}C)_{ij} : 1 \le i,j \le m\} \le x\right\}$$

for all $x$. Now

$$\begin{aligned}\mathbb{P}\left\{\bigvee\{(ABC)_{ij} : 1 \le i,j \le m\} \le x\right\} &= \mathbb{P}\left(\bigcap_{ijk\ell}\{A_{ij} + B_{jk} + C_{k\ell} \le x\}\right) \\ &= \prod_{h=1}^{n} \prod_{(j,k)\in\Gamma_h} \mathbb{P}\left\{B_{jk} \le \bigwedge_{i\ell}(x - A_{ij} - C_{k\ell})\right\}.\end{aligned}$$

On the other hand, writing $\hat{B}_h$ for the common value of $\tilde{B}_{jk}$, $(j,k) \in , _h$, we have

$$\mathbb{P}\left\{\bigvee\{(A\tilde{B}C)_{ij} : 1 \le i,j \le m\} \le x\right\} = \prod_{h=1}^{n} \mathbb{P}\left\{\hat{B}_h \le \bigwedge_{(j,k)\in\Gamma_h} \bigwedge_{i\ell}(x - A_{ij} - C_{k\ell})\right\}.$$

For $1 \le h \le n$ choose $(j^*, k^*) \in , _h$ such that

$$\bigwedge_{(j,k)\in\Gamma_h} \bigwedge_{i\ell}(x - A_{ij} - C_{k\ell}) = \bigwedge_{i\ell}(x - A_{ij^*} - C_{k^*\ell}).$$

Then

$$\prod_{(j,k)\in\Gamma_h} \mathbb{P}\left\{B_{jk} \le \bigwedge_{i\ell}(x - A_{ij} - C_{k\ell})\right\} \le \mathbb{P}\left\{B_{j^*k^*} \le \bigwedge_{i\ell}(x - A_{ij^*} - C_{k^*\ell})\right\}$$
$$= \mathbb{P}\left\{\hat{B}_h \le \bigwedge_{i\ell}(x - A_{ij^*} - C_{k^*\ell})\right\}$$
$$= \mathbb{P}\left\{\hat{B}_h \le \bigwedge_{(j,k)\in\Gamma_h}\bigwedge_{i\ell}(x - A_{ij} - C_{k\ell})\right\},$$

and the result follows. □

## 3. Asymptotics of the global maximum

We may use Lemma 1 to elucidate the asymptotic behavior of the maximum $G^*_{N,K}$, where $K$ is fixed. (In this section, we will take $K$ as fixed and drop it from the notation.) Asymptotically, $G^*_N$ is equal to

$$G^{**}_N := \bigvee_{\vec{x}\in S} \frac{1}{N} \sum_{i=0}^{N-K-1} F_i(x_i, \ldots, x_{i+K}).$$

Since the end does not wrap around, $G^{**}_N$ is the maximum entry of the max–plus product $A_0 A_1 \cdots A_{N-K-1}$, with the $A$'s i.i.d. matrices. The sequence of random variables $NG^{**}_N$ is subadditive with respect to the standard shift: that is,

$$\bigvee_{\vec{x},\vec{y}} A_0 \cdots A_{N-K-1}(\vec{x},\vec{y}) \le \bigvee_{\vec{x},\vec{y}} A_0 \cdots A_m(\vec{x},\vec{y}) + \bigvee_{\vec{x},\vec{y}} A_{m+1} \cdots A_{N-K-1}(\vec{x},\vec{y}).$$

By Kingman's subadditive ergodic theorem, this implies that $G^{**}_N$, hence also $G^*_N$, has an almost-sure limit $\lambda$. Since the $A$'s are i.i.d., Kolmogorov's zero–one law applies and $\lambda$ must be a constant, the max–plus top Lyapunov exponents. We refer the reader to [Bac92, JM94] for an indication of the literature on Lyapunov exponents for products of max–plus matrices.

This $\lambda$ satisfies the max–plus version of the Furstenberg–Kifer theorem, namely

**Lemma 5.** *Let $\mu$ be the distribution of an $n \times n$ $\mathbb{R}\cup\{-\infty\}$-valued matrix with no row identically $-\infty$, and $\mathcal{P}(\mu)$ the set of laws on $\mathbb{R}\cup\{-\infty\}$-valued $n$-vectors such that if $W$ has distribution $\mu$ and $Y$ has distribution $\nu \in \mathcal{P}(\mu)$, with $W$ and $Y$*

*independent,*

$$\Big(\bigvee_j (W(i,j)+Y(j)) - \bigvee_{k,j}(W(k,j)+Y(j))\Big)_{i=1}^n$$

*also has law* $\nu$*.*

*Then if* $W_0, W_1, \ldots$ *are i.i.d. max–plus matrices with common distribution* $\mu$*,*

$$\begin{aligned}\lim_{N\to\infty}\frac{1}{N}\bigvee_{i,j}(W_0\ldots W_{N-1})(i,j)\\ = \sup\Big\{\mathrm{E}_{\mu\otimes\nu}\Big[\bigvee_i\bigvee_j(W(i,j)+Y(j))\Big] \;:\; \nu\in\mathcal{P}(\mu)\Big\}.\end{aligned} \tag{3}$$

The proof is a straightforward adaptation of the proof of the analogue of the Furstenberg–Kifer theorem for i.i.d. non-negative matrices given in Section 4.5 of [HM95].

We now take $\mu$ to be the distribution of the matrices $A$ for our problem. The coordinates of $Y$ and $A$ are naturally indexed by $\{0,1\}^K$. The invariant distributions $\mathcal{P}(\mu)$ are significantly constrained. A measure on $\mathbb{R}^{2^K}$ will be called *coordinatewise stationary* if it is invariant under coordinate transformations of the form

$$(x_0,\ldots,x_i,\ldots,x_{K-1})\mapsto(x_0,\ldots,1-x_i,\ldots,x_{K-1}) \tag{4}$$

for any fixed $i$. If this property holds for $i\le j$, we will call the measure $j$-stationary.

**Lemma 6.** *The measures in* $\mathcal{P}(\mu)$ *are all coordinatewise stationary.*

*Proof.* Let $Y$ and $A$ be independent random variables, with $A$ having law $\mu$ and $Y$ law $\nu\in\mathcal{P}(\mu)$. We also define

$$Y'(\vec{x}) := \Big(\bigvee_{\vec{y}}(A(\vec{x},\vec{y})+Y(\vec{y})) - \bigvee_{\vec{z},\vec{y}}(A(\vec{z},\vec{y})+Y(\vec{y}))\Big).$$

Because of the definition of $A$, this reduces to defining

(5)

$$\begin{aligned}Y''(x_0,\ldots,x_{K-1}) &:= \bigvee\Big\{F(x_0,\ldots,x_{K-1},1)\;+\;Y(x_1,\ldots,x_{K-1},1)\;,\\ &\qquad F(x_0,\ldots,x_{K-1},0)\;+\;Y(x_1,\ldots,x_{K-1},0)\Big\},\text{ and}\\ Y'(x_0,\ldots,x_{K-1}) &:= Y''(x_0,\ldots,x_{K-1}) - \bigvee_{\vec{z}} Y''(\vec{z}).\end{aligned}$$

Suppose that $Y$ is $(j-1)$-stationary. We want to show that $Y''$ has the same distribution as $\widetilde{Y}''$, defined by

$$\widetilde{Y}''(x_0,\ldots,x_j,\ldots,x_{K-1}) = Y''(x_0,\ldots,1-x_j,\ldots,x_{K-1}).$$

Define

$$\widetilde{F}(x_0,\ldots,x_K) := F(x_0,\ldots,1-x_j,\ldots x_K) \text{ and}$$
$$\widetilde{Y}(x_0,\ldots,x_{K-1}) := Y(x_0,\ldots,1-x_j,\ldots,x_{K-1}).$$

Then $\widetilde{Y}''$ may be computed from (5) simply by substituting $\widetilde{F}$ for $F$ and $\widetilde{Y}$ for $Y$.

Observe now that $\widetilde{F}$ and $F$ have the same distribution. Also, since $Y$ is $(j-1)$-stationary, the joint distributions of $Y(x_1,\ldots,x_j,\ldots,x_K,x_K)$ and of $Y(x_1,\ldots,1-x_j,\ldots,x_K,x_K)$ are identical. Since $F$ and $Y$ are independent, it follows that $\widetilde{Y}''$ has the same distribution as $\widetilde{Y}$. It follows then that $Y'$ is also $j$-stationary, and $Y$ and $Y'$ have the same distribution. The result is that $Y$ is coordinatewise stationary. □

We now consider the case of $K=1$. The matrix $A$ is then $2\times 2$, with i.i.d. entries, while $Y$ is an $\mathbb{R}^2$-valued random variable, with identically distributed coordinates. Observe that one of the coordinates of $Y'$ is 0, while we may represent the other coordinate as $-\hat{Y}$; since the coordinates are identically distributed, it follows that $Y$ (hence $Y'$ as well) takes on the values $(-\hat{Y},0)$ and $(0,-\hat{Y})$ each with probability $\frac{1}{2}$.

Suppose that $F_i$ has density $f$ and distribution function $\mathcal{F}$, and suppose that $\hat{Y}$ has distribution $\mu$. The distribution function of $Y''$ is then

$$\begin{aligned}\mathbb{P}\{Y_1'' \le w,\ Y_2'' \le z\} &= \mathbb{E}\big[\mathcal{F}(w-Y_1)\mathcal{F}(w-Y_2)\mathcal{F}(z-Y_1)\mathcal{F}(z-Y_2)\big] \\ &= \mathcal{F}(w)\mathcal{F}(z)\,\mathbb{E}\big[\mathcal{F}(w+\hat{Y})\mathcal{F}(z+\hat{Y})\big] \\ &= \mathcal{F}(w)\mathcal{F}(z)\int \mathcal{F}(w+y)\mathcal{F}(z+y)\,\mu(dy).\end{aligned}$$

The random variable $Y''$ has a density

$$\begin{aligned} g(w,z) = f(w)f(z)\int \mathcal{F}(w+y)\mathcal{F}(z+y)\,\mu(dy) \\ + f(w)\mathcal{F}(z)\int \mathcal{F}(w+y)f(z+y)\,\mu(dy) \\ + \mathcal{F}(w)\mathcal{F}(z)\int f(w+y)f(z+y)\,\mu(dy) \\ + \mathcal{F}(w)f(z)\int f(w+y)\mathcal{F}(z+y)\,\mu(dy). \end{aligned}$$

Since $|Y_1'' - Y_2''|$ has the same distribution as $\hat{Y}$, we see that $\mu$ must have a density $\mu(dy) = h(y)\,dy$, with

$$h(u) = \int_{-\infty}^{\infty} \big[g(w, w+u) + g(w, w-u)\big]dw \tag{6}$$

for $u$ positive.

Consider now the case in which $F$ has the exponential distribution with expectation 1. Then, letting $\tilde{h}$ be the Laplace transform of $h$,

$$\begin{aligned} g(w,z) = e^{-w-z}\big(1 + 2\tilde{h}(1) + \tilde{h}(2)\big) + \big(e^{-2w-z} + e^{-2z-w}\big)\big(-2\tilde{h}(1) - 2\tilde{h}(2)\big) \\ + 4e^{-2w-2z}\tilde{h}(2)\,. \end{aligned}$$

Substituting into the (6) then yields

$$h(u) = e^{-u} + \left(\frac{2}{3}e^{-u} - \frac{4}{3}e^{-2u}\right)\tilde{h}(1) + \left(\frac{2}{3}e^{-2u} - \frac{1}{3}e^{-u}\right)\tilde{h}(2). \tag{7}$$

Multiplying by $e^{-u}$ or $e^{-2u}$ and integrating out, we conclude that

$$\begin{aligned} \tilde{h}(1) &= \frac{1}{2} - \frac{1}{9}\tilde{h}(1) + \frac{1}{18}\tilde{h}(2), \\ \tilde{h}(2) &= \frac{1}{3} - \frac{1}{9}\tilde{h}(1) + \frac{1}{18}\tilde{h}(2). \end{aligned}$$

These equations have the unique solution

$$\tilde{h}(1) = \frac{53}{114}, \qquad\qquad \tilde{h}(2) = \frac{34}{114}.$$

This means that there is a unique distribution in $\mathcal{P}(\mu)$, with density given by (7),

$$h(u) = \frac{23}{19}e^{-u} - \frac{8}{19}e^{-2u}.$$

The Lyapunov exponent is then given according to Lemma 5 by

$$\mathbb{E}\big[(F(0,0) - H) \vee (F(1,0) - H) \vee F(0,1) \vee F(1,1)\big],$$

where the $F$'s and $H$ are independent, with $H$ having density $h$, and the $F$'s having exponential distribution with expectation 1. If we let $S = F(0,0) \vee F(1,0)$, and $T = F(0,1) \vee F(1,1)$, then $S$ and $T$ are independent, with

$$\mathbb{P}\{S \leq x\} = \mathbb{P}\{T \leq x\} = (1 - e^{-x})^2, \quad x \geq 0,$$

and

$$\mathbb{P}\{S - H \leq x\} = 1 - \frac{53}{57} e^{-x} + \frac{17}{57} e^{-2x}, \quad x \geq 0.$$

Hence the Lyapunov exponent is

$$\begin{aligned} \mathbb{E}\big[(S - H) \vee T\big] &= \int_0^\infty \mathbb{P}\{(S - H) \vee T > x\}\, dx \\ &= \int_0^\infty 1 - \mathbb{P}\{(S - H) \leq x\}\, \mathbb{P}\{T \leq x\}\, dx \\ &= \frac{407}{228}. \end{aligned}$$

We may conclude that

$$\lim_{N \to \infty} G^*_{N,1} = \frac{407}{228} \approx 1.78509.$$

Observe that this is significantly larger than the expected height of a local maximum, which is 1.61651. We note that this example was worked out using somewhat different methods (and in a different context) as Proposition 4.3 in [JM94].

## 4. Counting local maxima

To compute the expected number of local maxima, we need to find the probability of any given point — for example, the point $\bar{0}_N$, the $N$-vector of all zeroes — being a local maximum. If $K = 0$ and $\mathcal{F}$ is a continuous distribution, then there is clearly only one local maximum. We therefore suppose for the remainder of this section that $K \geq 1$.

Letting $e^j_k$ be the $k$-vector with a single 1 in the $j$-th place (counting modulo $N$: if $j$ is not between 0 and $k$ modulo $N$, then $e^j_k := \bar{0}_k$), we see that $\bar{0}_N$ is a local maximum precisely when

$$G_{N,K}(\bar{0}_N) \geq G_{N,K}(e^j_N)$$

for all $0 \leq j \leq N-1$, which is equivalent to

$$\sum_{i=0}^{K} F_{j-K+i}(\bar{0}_{K+1}) \geq \sum_{i=0}^{K} F_{j-K+i}(e_{K+1}^{K+1-i}) \tag{8}$$

for all $0 \leq j \leq N-1$, where indices of the $F$'s are taken modulo $N$.

Define

$$X_i := F_i(\bar{0}_{K+1}),$$

$$Z_j := \sum_{i=0}^{K} F_{j-K+i}(e_{K+1}^{K+1-i}).$$

Then the condition for $\bar{0}_N$ to be a local maximum is

$$X_{j-K} + X_{j-K+1} + \cdots + X_j \geq Z_j,\ 0 \leq j \leq N-1. \tag{9}$$

The random variables $X_j$ and $Z_j$ are all independent. Conditioned on the values of the $X_j$'s, the probability of this event is thus

$$\mathcal{F}_*(X_0 + \cdots + X_K)\mathcal{F}_*(X_1 + \cdots + X_{K+1}) \cdots \mathcal{F}_*(X_{N-1} + X_0 + \cdots + X_{K-1}). \tag{10}$$

The probability of a point being a local maximum is then the expectation of (10) for $X_0, \ldots, X_{N-1}$ i.i.d. with distribution $\mathcal{F}$.

We wish to estimate this probability for large $N$. We begin by estimating

$$P_\nu := \mathbb{E}\left[\prod_{i=0}^{\nu(K+1)-1} \mathcal{F}_*(X_i + X_{i+1} + \cdots + X_{i+K})\right], \tag{11}$$

where $\nu = \lfloor N/(K+1) \rfloor - 1$. For $u, u' \in \mathbb{R}^K$ and $x \in \mathbb{R}$, define

$$\begin{aligned} H(u, x, u') := {} & \mathcal{F}_*(u_0 + \cdots + u_{K-1} + x)\mathcal{F}_*(u_1 + \cdots + u_{K-1} + x + u'_0) \\ & \cdots \mathcal{F}_*(u_{K-1} + x + u'_0 + \cdots + u'_{K-2})\mathcal{F}_*(x + u'_0 + \cdots + u'_{K-1}). \end{aligned}$$

Then, if we let $U_i := (X_{i(K+1)}, X_{i(K+1)+1}, \ldots, X_{i(K+1)+K-1})$,

$$\begin{aligned} P_\nu = {} & \mathbb{E}\big[H\left(U_0, X_K, U_1\right) H\left(U_1, X_{2K+1}, U_2\right) \\ & \cdots H\left(U_{\nu-1}, X_{(\nu-1)(K+1)+K}, U_\nu\right)\big]. \end{aligned}$$

The terms $X_K, X_{2K+1}, \ldots, X_{(\nu-1)(K+1)+K}$ may be integrated out, yielding

$$P_\nu = \mathbb{E}\big[\xi(U_0, U_1)\xi(U_1, U_2) \cdots \xi(U_{\nu-1}, U_\nu)\big], \tag{12}$$

where

$$\begin{aligned}\xi(u_0,\ldots,u_{K-1},u'_0,\ldots,u'_{K-1})&\\ := \int_{\mathbb{R}} \mathcal{F}_*(u_0+\cdots+u_{K-1}+z)&\mathcal{F}_*(u_1+\cdots+u_{K-1}+z+u'_0)\\ \cdots\mathcal{F}_*(z+u'_0+\cdots+u'_{K-1})\,&\mathcal{F}(dz).\end{aligned}$$

The kernel $\xi$ is bounded above by 1, and if the fitness variables are bounded below — that is, if $\mathcal{F}(c) = 0$ for some finite $c$ — then $\xi$ is bounded away from 0 when the domain is restricted to $[c,\infty)^K \times [c,\infty)^K$. Assume for the moment that this condition holds. By the obvious linearity, we may take $c = 0$ without loss of generality.

The distribution $\mathcal{F}$ defines an inner product on functions from $\mathbb{R}_+^K$ to $\mathbb{R}$ by

$$\langle f,g\rangle := \int_{\mathbb{R}_+^K} f(u)g(u)\,\mathcal{F}(du_0)\cdots\mathcal{F}(du_{K-1});$$

the corresponding space of square-integrable functions we will call simply $\mathcal{L}^2$. The kernel $\xi$ gives rise to an operator $\Xi$ on this space by

$$\Xi f(u) := \int_{\mathbb{R}_+^K} \xi(u,v)f(v)\,\mathcal{F}(dv_0)\cdots\mathcal{F}(dv_{K-1}) \text{ for } f\in\mathcal{L}^2. \tag{13}$$

The adjoint operator $\Xi^*$ is

$$\Xi^* g(v) := \int_{\mathbb{R}_+^K} \xi(u,v)g(u)\,\mathcal{F}(du_0)\cdots\mathcal{F}(du_{K-1}) \text{ for } g\in\mathcal{L}^2.$$

Then

$$P_\nu = \int_{\mathbb{R}_+^K} \Xi^{\nu-1}\mathbf{1}\,\mathcal{F}(du_0)\cdots\mathcal{F}(du_{K-1}),$$

where $\mathbf{1}$ is the constant function with value 1.

It is observed in Section 7 that the common spectral radius $\rho$ of $\Xi$ and $\Xi^*$ is an eigenvalue with multiplicity one for both operators. Let $\phi$ and $\psi$ be the corresponding eigenfunctions of $\Xi$ and $\Xi^*$ respectively, both chosen to be positive and bounded away from 0, and normalized to have norm 1. We note that $\psi(u_0,\ldots,u_{K-1}) = \phi(u_{K-1},\ldots,u_0)$. We may apply the Perron-Frobenius Theorem

in the version given here as Theorem 9 (in Section 7). This tells us that for any $g$ and $g^*$ in $\mathcal{L}^2(\mathbb{R}_+^K)$,

$$\lim_{\nu\to\infty} \rho^{-\nu}\, \mathbb{E}\big[g(U_0)\xi(U_0,U_1)\xi(U_1,U_2)\cdots\xi(U_{\nu-1},U_\nu)g^*(U_\nu)\big] = \frac{\langle g,\phi\rangle\langle g^*,\psi\rangle}{\langle \phi,\psi\rangle}.$$

The probability that we are looking for is

$$\begin{aligned}
&\mathbb{P}\big\{\bar{0}_N \text{ is a local max}\big\}\\
&\quad = \mathbb{E}\big[\xi(U_0,U_1)\xi(U_1,U_2)\cdots\xi(U_{\nu-1},U_\nu)\\
&\qquad \cdot \mathcal{F}_*(X_{\nu(K+1)+1}+\cdots+X_{(\nu+1)(K+1)})\\
&\qquad \cdots\mathcal{F}_*(X_{N-K}+\cdots+X_{N-1}+X_0)\cdots\mathcal{F}_*(X_{N-1}+X_0+\cdots+X_{K-1})\big].
\end{aligned}$$

Define a function

$$\begin{aligned}
&\gamma(u_0,\dots,u_{K-1};v_0,\dots,v_{K-1})\\
&:= \int\cdots\int \mathcal{F}_*(v_0+\cdots+v_{K-1}+x_{(\nu+1)(K+1)})\\
&\qquad \cdots\mathcal{F}_*(v_{N-K-\nu(K+1)}+\cdots+v_{K-1}+x_{\nu(K+1)+K}+\cdots+x_{N-1}+u_0)\\
&\qquad\quad \cdots\mathcal{F}_*(x_{N-1}+u_0+\cdots+u_{K-1})\,\mathcal{F}(dx_{\nu(K+1)+K})\cdots\mathcal{F}(dx_{N-1}).
\end{aligned} \tag{14}$$

This function is in $\mathcal{L}^2(\mathbb{R}_+^{m+2K})$ (where the inner product is again defined by the tensor power of the distribution $\mathcal{F}$, and $m = N-(\nu+1)(K+1)+1)$. Integrating out the $X$'s with indices between $(\nu+1)(K+1)-1$ and $N-1$, we find

$$\mathbb{P}\big\{\bar{0}_N \text{ is a local max}\big\} = \mathbb{E}\big[\xi(U_0,U_1)\xi(U_1,U_2)\cdots\xi(U_{\nu-1},U_\nu)\gamma(U_0,U_\nu)\big]. \tag{15}$$

Since we may approximate $\gamma$ by linear combinations of products of the form $\gamma_i(U_0)\gamma_i'(U_\nu)$, we may conclude that

$$\begin{aligned}
&\lim_{\nu\to\infty} \rho^{-\nu}\,\mathbb{P}\big\{\bar{0}_N \text{ is a local max}\big\} =\\
&\qquad \frac{1}{\langle\phi,\psi\rangle}\int_{\mathbb{R}_+^K}\int_{\mathbb{R}_+^K}\gamma(u,v)\psi(u)\phi(v)\,\mathcal{F}(du_0)\cdots\mathcal{F}(du_{K-1})\,\mathcal{F}(dv_0)\cdots\mathcal{F}(dv_{K-1}).
\end{aligned}$$

Observe that $\gamma$ is everywhere positive and less than 1. Since $\phi$ and $\psi$ are strictly positive, we have

**Theorem 7.** *Suppose that $K \geq 1$ and the fitness distribution satisfies $\mathcal{F}(0) = 0$. If $\Xi$ is the operator defined by (13) and $\rho$ its spectral radius, then*

$$1 \geq \lim_{n\to\infty} \rho^{-N/(K+1)}\, \mathbb{P}\big\{\bar{0}_N \text{ is a local max}\big\} > 0.$$

*Consequently,*

$$\lim_{N\to\infty} \mathbb{E}\big[\#\text{ local maxima}\big]^{1/N} = 2\rho^{1/(K+1)}.$$

The value of $\rho$ appears to depend on the distribution $\mathcal{F}$. We first compute it for the case of an exponential distribution with expectation 1. (Because any other exponential random variable is simply a constant multiple of this one, a different choice of expectation would lead to the same spectral radius.) In this case,

$$\mathcal{F}_*(z) = 1 - e^{-z} \sum_{i=0}^{K} \frac{z^i}{i!}.$$

We define the Laplace transform of the kernel $\xi$ with respect to the first variable by

$$\begin{aligned}(16)\qquad \tilde{\xi}(\alpha, v) &:= \int_{\mathbb{R}_+^K} e^{-\alpha\cdot u}\xi(u, v)du_0 \cdots du_{K-1} \\ &= \int_{\mathbb{R}_+^K}\int_{\mathbb{R}_+} e^{-\alpha\cdot u - z}\mathcal{F}_*(u_0 + \cdots + u_{K-1} + z) \\ &\qquad\qquad \cdots \mathcal{F}_*(z + v_0 + \cdots + v_{K-1})\,\mathcal{F}(dz)du_0 \cdots du_{K-1},\end{aligned}$$

and define the Laplace transform of the Perron-Frobenius eigenfunction $\phi$ by

$$\tilde{\phi}(\alpha_0, \ldots, \alpha_{K-1}) := \int_{\mathbb{R}_+^K} \phi(u_0, \ldots, u_{K-1})e^{-\alpha\cdot u}du_0 \cdots du_{K-1}.$$

Then

$$(17)\qquad \rho\tilde{\phi}(\alpha) = \int_{\mathbb{R}_+^K} \tilde{\xi}(\alpha, v)\phi(v)e^{-v_0 - v_1 - \cdots - v_{K-1}}dv_0 \cdots dv_{K-1}.$$

Consider the case $K = 1$. We get

$$\begin{aligned}\tilde{\xi}(\alpha_0; v_0) = \int_0^\infty \int_0^\infty & e^{-\alpha_0 w_0 - z}\Big(1 - e^{-w_0 - z}(1 + w_0 + z)\Big) \\ &\times \Big(1 - e^{-v_0 - z}(1 + v_0 + z)\Big)dw_0\, dz.\end{aligned}$$

Write (with $\alpha = \alpha_0$ and $v = v_0$),

$$\tilde{\xi}(\alpha; v) = \frac{1}{108\alpha(1+\alpha)^2}\Big[(108 + 81\alpha + 27\alpha^2) - (81 + 46\alpha + 13\alpha^2)e^{-v} + (54 + 24\alpha + 6\alpha^2)ve^{-v}\Big].$$

Plugging this into (17) yields

$$\rho\tilde{\phi}(\alpha) = \frac{1}{108\alpha(1+\alpha)^2}\Big[(108 + 81\alpha + 27\alpha^2)\tilde{\phi}(1) - (81 + 46\alpha + 13\alpha^2)\tilde{\phi}(2) + (54 + 24\alpha + 6\alpha^2)\tilde{\phi}'(2)\Big].$$

The vector $[\tilde{\phi}(1)\ \tilde{\phi}(2)\ \tilde{\phi}'(2)]^\top$ is then a right-eigenvector of the $3 \times 3$ matrix

$$\frac{1}{3888}\begin{pmatrix} 1944 & -1260 & 756 \\ 756 & -450 & 252 \\ -504 & 329 & -198 \end{pmatrix}$$

The largest eigenvalue is $\rho \approx .316611$, which means, by Theorem 7, that

$$\lim_{N\to\infty} \mathbb{P}\{\bar{0}_N \text{ is a local maximum}\}^{1/N} \approx \sqrt{.316611} \approx .562682.$$

This means, in turn, that the expected number of local maxima grows approximately as $1.12536^N$.

In principle, the same method could be applied to any fixed value of $K$. In practice, the computations quickly become unmanageable. For example, when $K = 2$, we get, in place of the above matrix, an unprintable $22 \times 22$ matrix, and the principle eigenvector corresponds to the Laplace transform $\tilde{\phi}$ and various of its mixed derivatives up to total order 4 evaluated at pairs of arguments taken from $\{1, 2, 3\}$. The principle eigenvalue is approximately .228558, with cube root .611409. Thus, for $K = 2$, the expected number of local maxima grows approximately as $1.22282^N$.

A similar reduction via Laplace transforms to a finite dimensional eigenvalue problem occurs when $\mathcal{F}$ is any gamma distribution with integer shape parameter. For example, when the shape parameter is 2 (so that $\mathcal{F}$ is the distribution of the sum of two i.i.d. exponentials) and $K = 1$, the expected number of local maxima grows approximately as $1.12915^N$. In particular, this growth rate differs (albeit slightly)

from the $K = 1$ growth rate for exponentially distributed fitnesses – bolstering the belief that the growth rate depends on the details of $\mathcal{F}$ in a rather complex manner.

## 5. Heights of the local maxima

For any point in $\{0,1\}^N$ — in particular, for the point $\bar{0}_N$ — the expected value of $G_{N,K}$ is the same as the expected value of $F_i$. If the point is known to be a local maximum, on the other hand, this should increase the expected value of $G_{N,K}$. We have

$$\begin{aligned}&\mathbb{E}\big[G_{N,K}(0,0,\ldots,0)\,\big|\,\bar{0}_N \text{ is a local max}\big]\\ &\qquad = \frac{\mathbb{E}\big[F_{[N/2]}(0,\ldots,0)\mathbf{1}\{\bar{0}_N \text{ is a local max}\}\big]}{\mathbb{P}\big\{\bar{0}_N \text{ is a local max}\big\}}.\end{aligned} \tag{18}$$

Of course, the expectation would be the same for any $F_i$ in place of $F_{[N/2]}$; we choose the coordinate in the middle merely to keep it away from the messy indexing behavior that occurs when we arbitrarily break the loop of dependencies between coordinate $N-1$ and 0.

Recall our notation $X_i = F_i(0,\ldots,0)$. As usual, we consider the indices modulo $N$. The $X_i$'s are independent random variables with distribution function $\mathcal{F}$. Then the numerator in (18) is

$$\mathbb{E}\big[X_{[N/2]}\mathcal{F}_*(X_0+\cdots+X_K)\cdots\mathcal{F}_*(X_{N-1}+\cdots+X_{K-1})\big],$$

and the denominator is the same without the factor of $X_{[N/2]}$.

From Section 4 we know that when $K \geq 1$ and $\mathcal{F}(0) = 0$ (so that the $X_i$ are non-negative)

$$\begin{aligned}&\lim_{N\to\infty} \rho^{-N/(K+1)}\,\mathbb{P}\big\{\bar{0}_N \text{ is a local max}\big\}\\ &\qquad = \frac{1}{\langle\phi,\psi\rangle}\int_{\mathbb{R}_+^K}\int_{\mathbb{R}_+^K}\gamma(u,v)\psi(u)\phi(v)\,\mathcal{F}(du_0)\cdots\mathcal{F}(du_{K-1})\,\mathcal{F}(dv_0)\cdots\mathcal{F}(dv_{K-1}).\end{aligned}$$

Furthermore, if we slip a third function $h$ into the middle of the product — so $h(U_\mu)$, where $\mu$ and $\nu - \mu$ both go to infinity — we get

$$\begin{aligned}\lim_{\mu\to\infty,\ \nu-\mu\to\infty} \rho^{-\nu}\,\mathbb{E}\big[g(U_0)\xi(U_0,U_1)\cdots\xi(U_{\mu-1},U_\mu)h(U_\mu)\\ \cdot\xi(U_\mu,U_{\mu+1})\cdots\xi(U_{\nu-1},U_\nu)g^*(U_\nu)\big]\\ = \frac{\langle g,\phi\rangle\langle g^*,\psi\rangle\langle h\phi,\psi\rangle}{\langle\phi,\psi\rangle^2}.\end{aligned}$$

It follows that

$$\begin{aligned}&\lim_{N\to\infty} \rho^{-N/(K+1)}\,\mathbb{E}\big[X_{[N/2]}\mathcal{F}_*(X_0+\cdots+X_K)\cdots\mathcal{F}_*(X_{N-1}+\cdots+X_{K-1})\big\}\\ &\quad= \frac{1}{\langle\phi,\psi\rangle^2}\int_{\mathbb{R}_+^K}\psi(u)\phi(u)u_0\,\mathcal{F}(du_0)\cdots\mathcal{F}(du_{K-1})\\ &\qquad\cdot\int_{\mathbb{R}_+^K}\int_{\mathbb{R}_+^K}\gamma(u,v)\psi(u)\phi(v)\,\mathcal{F}(du_0)\cdots\mathcal{F}(du_{K-1})\,\mathcal{F}(dv_0)\cdots\mathcal{F}(dv_{K-1}).\end{aligned}$$

The conclusion is the following.

**Theorem 8.** *Suppose that $\mathcal{F}(0) = 0$. The expected height of the fitness function at a point, conditioned on the point being a local maximum, converges to*

$$\frac{\int_{\mathbb{R}_+^K}\psi(u)\phi(u)u_0\,\mathcal{F}(du_0)\cdots\mathcal{F}(du_{K-1})}{\int_{\mathbb{R}_+^K}\psi(u)\phi(u)\,\mathcal{F}(du_0)\cdots\mathcal{F}(du_{K-1})},$$

*as $N\to\infty$, where $\phi$ (respectively, $\psi$) is the Perron–Frobenius eigenfunction for the operator $\Xi$ (resp. $\Xi^*$), defined in* (13).

Consider the case when $\mathcal{F}$ is the exponential distribution with expectation 1. The expected height of an ordinary point is 1 for all $K$. When $K = 0$ the expected height of a point conditioned on it being a local maximum is readily seen to be the expectation of the maximum of two independent exponential random variables, each with expectation 1, and hence this conditional expected value is $3/2 = 1.5$. When $K = 1$ we have already computed $\sqrt{\rho} \approx .562682$ and the Laplace transform of the corresponding eigenfunction. By inverting the Laplace transform we see that the eigenfunction is

$$\phi(v) = \psi(v) \approx 2.18043 - 1.45895e^{-v} - 0.896269ve^{-v}.$$

The integral of $\phi\psi$ against $e^{-v}$ is about 1.65561, while the integral against $ve^{-v}$ is about 2.67631. Thus, the expected height of a local maximum converges to 1.61651.

In the case $K = 2$, the expected height of a local maximum is, by similar computations, asymptotic to 1.86367. This increase with $K$ of the expected height of local maxima is noteworthy, inasmuch as Kauffman found that the height of the local maximum attained by hill-climbing from a random starting point seemed to increase in $K$, at least for the first few values of $K$. He explained this by saying that higher peaks have larger basins of attraction.

We note that if $\mathcal{F}$ is the distribution of the sum of two i.i.d. exponential random variables with common expectation 1, then the expected height of a local maximum converges to 2.88039 for $K = 1$. The expected height for $K = 0$ is $11/4 = 2.75$.

## 6. An alternative representation

We want to give an alternative expression for the probability that a point, say $\bar{0}_N$ is a local maximum in the case when the fitnesses are exponential random variables. (The probability is obviously invariant under rescaling of the fitnesses, so we can take the fitnesses to have mean 1.)

This probability that $\bar{0}_N$ is a local maximum is (recalling the notation $X_i := F_i(\bar{0}_{K+1})$)

$$
\begin{aligned}
&\mathbb{E}\left[\prod_{j=0}^{N-1} \mathcal{F}_*(X_j + \cdots + X_{j+K})\right] \\
&= \mathbb{E}\left[\prod_{j=0}^{N-1} \left\{ \sum_{k=K+1}^{\infty} \exp(-(X_j + \cdots + X_{j+K})) \sum_{\ell_0 + \cdots + \ell_K = k} \frac{X_j^{\ell_0}}{\ell_0!} \cdots \frac{X_{j+K}^{\ell_K}}{\ell_K!} \right\}\right] \\
&= \mathbb{E}\Bigg[ \sum_{k_0=K+1}^{\infty} \cdots \sum_{k_{N-1}=K+1}^{\infty} \sum_{\ell_{0,0}+\cdots+\ell_{0,K}=k_0} \cdots \sum_{\ell_{N-1,0}+\cdots+\ell_{N-1,K}=k_{N-1}} \\
&\exp(-(X_0 + \cdots + X_K)) \frac{X_0^{\ell_{0,0}} \cdots X_K^{\ell_{0,K}}}{\ell_{0,0}! \cdots \ell_{0,K}!} \times \cdots \\
&\cdots \times \exp(-(X_{N-1} + \cdots + X_{N-1+K})) \frac{X_{N-1}^{\ell_{N-1,0}} \cdots X_{N-1+K}^{\ell_{N-1,K}}}{\ell_{N-1,0}! \cdots \ell_{N-1,K}!} \Bigg] \\
&= \mathbb{E}\Bigg[ \sum_{k_0=K+1}^{\infty} \cdots \sum_{k_{N-1}=K+1}^{\infty} \sum_{\ell_{0,0}+\cdots+\ell_{0,K}=k_0} \cdots \sum_{\ell_{N-1,0}+\cdots+\ell_{N-1,K}=k_{N-1}} \\
&\exp(-(K+1)X_0) \frac{X_0^{\ell_{-K,K}+\ell_{-(K-1),K-1}+\cdots+\ell_{0,0}}}{\ell_{-K,K}!\ell_{-(K-1),K-1}! \cdots \ell_{0,0}!} \\
&\times \exp(-(K+1)X_1) \frac{X_1^{\ell_{1-K,K}+\ell_{1-(K-1),K-1}+\cdots+\ell_{1,0}}}{\ell_{1-K,K}!\ell_{1-(K-1),K-1}! \cdots \ell_{1,0}!} \times \cdots \\
&\cdots \times \exp(-(K+1)X_{N-1}) \frac{X_{N-1}^{\ell_{N-1-K,K}+\ell_{N-1-(K-1),K-1}+\cdots+\ell_{N-1,0}}}{\ell_{N-1-K,K}!\ell_{N-1-(K-1),K-1}! \cdots \ell_{N-1,0}!} \Bigg].
\end{aligned}
$$

Because

$$
\int_0^\infty \exp(-(K+1)x) x^\ell \exp(-x)\, dx = \left(\frac{1}{K+2}\right)^{\ell+1} \ell!,
$$

the quantity we seek is

$$
\begin{aligned}
&\sum \frac{1}{K+2}\left(\frac{K+1}{K+2}\right)^{\ell_{-K,K}+\ell_{-(K-1),K-1}+\cdots+\ell_{0,0}} \\
&\times \frac{(\ell_{-K,K}+\cdots+\ell_{0,0})!}{\ell_{-K,K}!\cdots\ell_{0,0}!}\left(\frac{1}{K+1}\right)^{\ell_{-K,K}}\cdots\left(\frac{1}{K+1}\right)^{\ell_{0,0}} \\
&\times \frac{1}{K+2}\left(\frac{K+1}{K+2}\right)^{\ell_{1-K,K}+\ell_{1-(K-1),K-1}+\cdots+\ell_{1,0}} \\
&\times \frac{(\ell_{1-K,K}+\cdots+\ell_{1,0})!}{\ell_{1-K,K}!\cdots\ell_{1,0}!}\left(\frac{1}{K+1}\right)^{\ell_{1-K,K}}\cdots\left(\frac{1}{K+1}\right)^{\ell_{1,0}}\times \\
&\cdots \\
&\times \frac{1}{K+2}\left(\frac{K+1}{K+2}\right)^{\ell_{N-1-K,K}+\ell_{N-1-(K-1),K-1}+\cdots+\ell_{N-1,0}} \\
&\times \frac{(\ell_{N-1-K,K}+\cdots+\ell_{N-1,0})!}{\ell_{N-1-K,K}!\cdots\ell_{N-1,0}!}\left(\frac{1}{K+1}\right)^{\ell_{N-1-K,K}}\cdots\left(\frac{1}{K+1}\right)^{\ell_{N-1,0}},
\end{aligned}
$$

where the sum is over all non-negative integers $\ell_{0,0},\ldots,\ell_{N-1,K}$ such that

$$
\begin{aligned}
&\ell_{0,0}+\ell_{0,1}+\cdots+\ell_{0,K}\geq K+1,\\
&\ell_{1,0}+\ell_{1,1}+\cdots+\ell_{1,K}\geq K+1,\\
&\ldots,\\
&\ell_{N-1,0}+\ell_{N-1,1}+\cdots+\ell_{N-1,K}\geq K+1.
\end{aligned}
$$

This last expression has a simple probabilistic interpretation. Let $S_0,\ldots,S_{N-1}$ be independent, identically distributed, non-negative, integer-valued random variables with

$$
\mathbb{P}\{S_j=s\}=\frac{1}{K+2}\left(\frac{K+1}{K+2}\right)^s,\quad s=0,1,2,\ldots
$$

That is, $S_i$ has the distribution of the number of failures before the first success in a sequence of i.i.d. Bernoulli trials with success probability $\frac{1}{K+2}$. Given $S_0,\ldots,S_{N-1}$, let the random vectors $T_j=(T_{j-K,K},T_{j-(K-1),K-1},\ldots,T_{j,0})$, $0\leq j\leq N-1$, be conditionally independent, with the conditional distribution of $T_j$

being multinomial$(S_j; \frac{1}{K+1}, \frac{1}{K+1}, \dots, \frac{1}{K+1})$. Equivalently,

$$\begin{aligned}&\mathbb{E}\left[\prod_{j=0}^{N-1} \tau_{j-K,K}^{T_{j-K,K}} \tau_{j-(K-1),K-1}^{T_{j-(K-1),K-1}} \cdots \tau_{j,0}^{T_{j,0}}\right] \\ &\quad = \prod_{j=0}^{N-1} \left(K+2-\left(\tau_{j-K,K}+\tau_{j-(K-1),K-1}+\cdots+\tau_{j,0}\right)\right)^{-1}\end{aligned}$$

Set $U_k := T_{k,0} + T_{k,1} + \cdots + T_{k,K}$, $0 \le k \le N-1$, so that

$$\mathbb{E}\left[\prod_{k=0}^{N-1} v_k^{U_k}\right] = \prod_{j=0}^{N-1} \left(K+2-\left(v_{j-K}+v_{j-(K-1)}+\cdots+v_j\right)\right)^{-1}.$$

It is easy to see (for example, by using the above probability generating functions) that each $U_k$ is distributed as the number of failures before the $(K+1)^{\text{st}}$ success in i.i.d. Bernoulli trials with success probability $\frac{1}{2}$. In particular, $\mathbb{E}[U_k] = K+1$. Of course, the $U_k$ are not independent, but $(U_0, \dots U_{N-1})$ is a stationary process on the group of integers modulo $N$. Then

$$\mathbb{P}\left\{\bar{0}_N \text{ is a local maximum}\right\} = \mathbb{P}\{U_k \ge K+1,\ 0 \le k \le N-1\}.$$

It is apparent from the probability generating function that the collection of random vectors $(T_{j-(K-1),K-1}, T_{j-(K-2),K-2}, \dots, T_{j,0})$, $0 \le j \le N-1$, has the same joint distribution as the analogue of the $T_j$'s for the $N(K-1)$ model. Therefore, if we set $\tilde{U}_k := T_{k,0} + T_{k,1} + \cdots + T_{k,K-1}$, $0 \le k \le N-1$, then the probability that $\bar{0}_N$ is a local maximum for the $N(K-1)$ model is

$$\mathbb{P}\{\tilde{U}_k \ge K,\ 0 \le k \le N-1\}.$$

The $N \to \infty$ asymptotics for $K = 0, 1, 2$ obtained in Section 4 suggest that for fixed $N$ the probability $\bar{0}_N$ is a local maximum for the $NK$ model might increase with $K$ (at least for exponentially distributed fitnesses). The "coupling" of the $NK$ and $N(K-1)$ models we have just described suggests a route to verifying this conjecture, but we are unable to supply a proof.

## 7. Some Perron–Frobenius Theory

Suppose that on some probability space $(\Sigma, \mathcal{A}, \mu)$ we have an $\mathcal{A} \times \mathcal{A}$–measurable kernel $\xi : \Sigma \times \Sigma \to \mathbb{R}$ that satisfies $0 < c \leq \xi \leq C < \infty$ for constants $c, C$. Define a compact operator $\Xi$ on the complex Hilbert space $L^2(\mu)$ by

$$\Xi g(x) := \int_\Sigma \xi(x, y) g(y)\, \mu(dy),$$

so that $\Xi$ has adjoint $\Xi^*$ given by

$$\Xi^* f(y) := \int_\Sigma f(x) \xi(x, y)\, \mu(dx).$$

As compact operators, $\Xi$ and $\Xi^*$ have a common spectrum that is discrete outside any neighborhood of 0. All nonzero elements of the spectrum are eigenvalues with finite multiplicity. Write $\rho$ for the common spectral radius of $\Xi$ and $\Xi^*$. That is, $\rho$ is the modulus of the largest eigenvalue. Equivalently,

$$\rho = \lim_n \|\Xi^n\|^{\frac{1}{n}} = \lim_n \|\Xi^{*n}\|^{\frac{1}{n}}.$$

Clearly, $c \leq \rho \leq C$.

The following result can be proved along the same lines as the classical Perron–Frobenius theorem for positive matrices (see, for example, [HJ85]). It is probable that this result exists in the literature, but we have been unable to find a suitable reference.

**Theorem 9.** *The spectral radius $\rho$ is an eigenvalue of $\Xi$ and $\Xi^*$ and is the unique eigenvalue with modulus $\rho$. Moreover, $\rho$ is simple for both $\Xi$ and $\Xi^*$. Let $\phi$ and $\psi$ be normalized eigenfunctions of $\Xi$ and $\Xi^*$ for the eigenvalue $\rho$ (so that $\phi$ and $\psi$ are unique up to constants of modulus 1). It is possible to choose constants so that $\phi \geq 0$ and $\psi \geq 0$, $\mu$–a.e., in which case $0 < \operatorname{ess\,inf} \phi \leq \operatorname{ess\,sup} \phi < \infty$ and $0 < \operatorname{ess\,inf} \psi \leq \operatorname{ess\,sup} \psi < \infty$. Finally, $\lim_n \|\rho^{-n} \Xi^n - \Lambda\| = \lim_n \|\rho^{-n} \Xi^{*n} - \Lambda^*\| = 0$, where $\Lambda$ is the rank one operator defined by*

$$\Lambda g(x) := \frac{\phi(x) \langle \psi, g \rangle}{\langle \phi, \psi \rangle}.$$

**Acknowledgement:** We thank the referee for a number of helpful comments.

## References

[Bac92] François Baccelli. Ergodic theory of stochastic Petri networks. *Ann. Probab.*, 20(1):375–396, 1992.

[DL01] Richard Durrett and Vlada Limic. Rigorous results for the $NK$ model. Preprint. Department of Mathematics, Cornell University, 2001.

[HJ85] Roger A. Horn and Charles R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, 1985.

[HM95] Göran Högnäs and Arunava Mukherjea. *Probability measures on semigroups*. Plenum Press, New York, 1995. Convolution products, random walks, and random matrices.

[JM94] Alain Jean-Marie. Analytical computation of Lyapunov exponents in stochastic event graphs. In *Performance evaluation of parallel and distributed systems: solution methods, Part 2 (Torino, 1993)*, pages 309–341. Math. Centrum Centrum Wisk. Inform., Amsterdam, 1994.

[Kah85] Jean-Pierre Kahane. *Some random series of functions*. Cambridge University Press, Cambridge, second edition, 1985.

[Kau93] Stuart Kauffman. *The origins of order*. Oxford University Press, 1993.

[KL87] S.A. Kauffman and S.A. Levin. Towards a general theory of adaptive walks on rugged landscapes. *J. Theor. Biol*, 128:11–45, 1987.

[SBTG99] Daniel Solow, Apostolos Burnetas, Ming-Chi Tsai, and Neil S. Greenspan. Understanding and attenuating the complexity catastrophe in Kauffman's NK model of genome evolution. *Complexity*, 5(1):53–66, 1999.

[SJ94] Dietrich Stauffer and Naeem Jan. Size effects in Kauffman type evolution for rugged fitness landscapes. *Journal of Theor. Biol.*, 168:211–218, 1994.

[Wei91] Edward D. Weinberger. Local properties of Kauffman's N-k model: A tunably rugged energy landscape. *Physical Review A*, 44(10):6399–6413, November 15 1991.

[Wil98] Claus O. Wilke. Evolution in time-dependent fitness landscapes. Technical Report 98-09, Institut für Neuroinformatik, Ruhr-Universität Bochum, November 13 1998. arXiv:physics/9811021.

[Wri32] Sewall Wright. The roles of mutation, inbreeding, crossbreeding, and selection in evolution. In Donald Jones, editor, *Proceedings of the VI international congress of Genetics*, volume 1, pages 356–66, 1932. Reprinted in *Evolution*, ed. Mark Ridley, Oxford University Press, 1997.

Department of Statistics #3860, University of California, 367 Evans Hall, Berkeley, CA 94720-3860, U.S.A.

*E-mail address*: evans@stat.berkeley.edu

Department of Demography #2120, University of California, 2232 Piedmont Avenue, Berkeley, CA 94720-2120, U.S.A.

*E-mail address*: dstein@demog.berkeley.edu