On the Expected Performance of Systems with Complex Interactions Among Components

Daniel Solow[*]
Apostolos Burnetas[†]
Department of Operations,
Case Western Reserve University,
Cleveland, OH 44106

Ming-chi Tsai
Department of Management Science,
Chinese Military Academy,
Taiwan, 830, R.O.C.

Neil S. Greenspan[‡]
Institute of Pathology,
Case Western Reserve University,
Cleveland, OH 44106

A class of combinatorial optimization models is presented for studying certain systems that arise in biology, physics, business, and elsewhere. These systems consist of a finite number of parts. For each part, it is necessary to choose one of several interchangeable components so as to maximize a performance measure of the resulting system that depends on how the chosen parts interact with each other. Probabilistic analysis and computer simulations provide insight into some factors that affect the expected performance of such systems. These models provide the ability to control the interactions among the components and to study the effect of replacing a single component in the system.

1. Introduction

A class of combinatorial optimization models is presented for studying systems composed of a finite number of parts. For each part, it is necessary to choose one of several interchangeable components that then interact with each other in complex ways that often cannot be measured. One objective in designing such a system is to choose, for each part, one of the available components in such a way that the resulting system is the best, according to a specific measure of performance.

[*]Electronic mail address: dxs8@po.cwru.edu.
[†]Electronic mail address: atb4@po.cwru.edu.
[‡]Electronic mail address: nsg@po.cwru.edu.

Complex Systems, **12** (2000) 423–456; © 2000 Complex Systems Publications, Inc.

One such example arises in the study of the maturation of the immune response [1], in which an antibody (the "system") is a collection of amino acid sites (the "parts") in the V region. At each site is one of 20 amino acids (the "components"). The affinity of an antibody for a particular antigen, which depends on how the chosen amino acids interact with each other, is a measure of the fitness (the "performance") of the resulting antibody.

Another example arises in the study of the evolution of genomes [2]. Here, a genome (the system) consists of a number of gene positions, or *loci* (the parts). At each locus, evolution selects one of several versions of a gene, that is, the *alleles* (the components). The *fitness* (the performance) of the resulting genome is based on how the selected genes interact with each other. It is assumed that genomes evolve so that allelic substitutions at each locus tend to maximize genomic fitness.

A third example of such a system arises in physics in the study of spin glasses [3]. Here, the system consists of a number of contiguous atoms (the parts). For each atom, it is possible to select a spin up or a spin down (the components). The total energy (the performance) of the atoms depends on how the selected spins interact with each other. The objective is to determine the spin of each atom so that the resulting ensemble has the least total energy.

Such systems also arise in the study of business organizations. For example, in [4] a special case of the model proposed here is used to study the expected performance of the team-replacement problem in an organization. In this setting, a *team* (the system) consists of a number of *job positions* (the parts). For each job position, it is possible to select one of a number of *qualified individuals* (the components). The *effectiveness* (the performance) of such a team is based on how the selected individuals interact with each other and the objective is to determine who to select for each position so that the resulting team is most effective.

A paradigm in the context of organizational change is developed in [5]. In that article, the system represents an organization having a number of attributes (the parts). For each attribute, it is possible to choose one of two alternative forms (the components). The author examines the effect of the amount of interaction among these attributes on the overall effectiveness (the performance) of the organization.

The relevance of complex systems to business organizations is further highlighted by the fact that an entire special issue of *Organization Science* [volume 10(3), 1999] was devoted to the "Application of Complexity Theory to Organization Science." Although that journal focuses more on the applications, rather than the mathematics, of models of complex systems to organizations, they indicate the need to develop more realistic models. The work here is an attempt to do just that and to obtain useful analytical and simulation results that provide insights into the expected performance of general complex systems, regardless

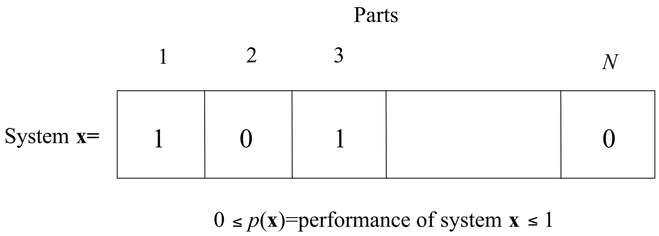

Figure 1. A system $\mathbf{x}$ as a binary N-vector.

of where these systems arise. A general class of models for studying this type of problem is developed in section 2. Some computational complexity results, polynomially-solvable special cases, and heuristics are then presented in section 3 for finding a system with relatively good performance. In section 4, conditions are developed under which it is possible to obtain analytical results on the expected performance of the best system found by a local improvement algorithm. The results are supported throughout by simulations. A summary and directions for future research conclude this work.

2. A class of combinatorial optimization models

A class of combinatorial optimization models for studying the type of problem presented in section 1 is developed in this section. A simplified version is described first so that the key features are easier to understand. A more general model is presented in section 2.2.

2.1 The basic model

Consider a system consisting of N parts, in which there are only two components available for each part. In this case, the system is represented mathematically as a binary N-vector, $\mathbf{x} = (x_1, \ldots, x_N)$, in which $x_i = 0$ means that one of the two components is chosen for the part in position i and $x_i = 1$ means that the other component is chosen for that part (see Figure 1). Each of the 2^N possible combinations of components in all N positions is referred to hereafter as a *system*. Geometrically, each of the 2^N systems corresponds to a corner point of the N-dimensional unit cube.

Each fixed choice of components for the N parts results in a system $\mathbf{x}$ whose relative performance is modeled as a real number $p(\mathbf{x})$ between 0 and 1. A value close to 1 indicates a system with relatively good performance and a value close to 0 indicates a system with relatively poor performance. In general, computing a value in an actual application is challenging because it may not be clear what is meant by *performance* nor how to compute a number in a realistic way that takes into account the complex interactions among the various components.

For systems whose performance is not measurable, a naïve approach is to generate, for each system, a uniform random number between 0 and 1 that represents the performance of that system. Although not realistic, some fundamental insights can still be obtained from such a model. Another more realistic approach is described next.

2.1.1 Computing the performance of a system using the *NK* model

In studying the maturation of the immune response, Weinberger [1] developed the *NK* model, which was subsequently applied to the evolution of genomes by Kauffman [2]. The approach to computing the fitness of a genome in the *NK* model is described now in the context of evaluating the performance of a system. It is assumed that each component chosen for part i contributes an amount $p_i(\mathbf{x})$ to the overall performance of the system $\mathbf{x}$. The performance $p(\mathbf{x})$ of the whole system $\mathbf{x}$ is then taken to be the average of these individual performance contributions:

$$p(\mathbf{x}) = \frac{\sum_{i=1}^{N} p_i(\mathbf{x})}{N}. \tag{1}$$

The remaining issue is how the performance contribution of the component chosen for part i is determined. In that regard, the *NK* model is designed to incorporate interaction among the chosen components through the use of an integer parameter K. The value of K $(0 \leq K \leq N-1)$ represents the number of other components that affect the performance contribution of the component chosen for part i. Thus, $K = 0$ indicates that the contribution of each component depends on no other components and $K = N - 1$ indicates that this contribution depends on all remaining $N - 1$ chosen components. More specifically, in the *NK* model, the contribution $p_i(\mathbf{x})$ of the component chosen for part i to the overall performance of the system $\mathbf{x}$ depends on the component in part i and on the components chosen for K other parts (e.g., the $K/2$ parts on either side of part i, wrapping around, if necessary). There are 2^{K+1} possible combinations for the components in these $K+1$ positions, so Kauffman defines the value of $p_i(\mathbf{x})$ to be one of 2^{K+1} uniform [0, 1] random numbers—the one that corresponds to the combination of the components for part i and the K components that affect component i. Given values for N, K, and the N tables of 2^{K+1} uniform [0, 1] random numbers, the collection of all 2^N binary vectors, together with their performance values, as defined by equation (1), constitute the *NK model*.

2.1.2 A heuristic for finding a good system in the *NK* model

One (or more) of these 2^N systems is the most effective—the one whose performance measure is closest to 1. Finding such a system in the *NK* model is not practical due to the *NP*-completeness result presented in section 3. Although Kauffman was unaware of this result, in the biological setting of genome evolution, he suggested a heuristic based on mutation. In the context of complex systems, that heuristic proceeds as follows. Starting with an initial system $\mathbf{x}$ a new system $\mathbf{x}'$ is created by considering what happens if a single component is replaced with the other available component for that part, resulting in what is called here a *one-replacement neighbor of* $\mathbf{x}$. The new system is retained only if $\mathbf{x}'$ has better performance than $\mathbf{x}$, that is, if $p(\mathbf{x}') > p(\mathbf{x})$. This *replacement process* results in a sequence of systems, each with better performance than its predecessor system, until obtaining a system whose performance is greater than or equal to that of all its one-replacement neighbors. This final system is a local maximum that is referred to here as a *locally-stable system*.

One question Kauffman set out to answer in this framework was how the values of N and K affect the average performance of the final system obtained from the replacement process. For the extreme values of $K = 0$ and $K = N - 1$, Kauffman was able to obtain analytical results that are then supported by simulations. For the case $K = 0$, the contribution of each component to the overall performance of the system depends only on that component. Starting with any initial system, a system with maximum performance is obtained by a replacement process that successively chooses the best component for each part, as determined by the random-number table for that part. The expected performance of the system thus obtained is shown mathematically to be 2/3 and the expected number of replacements needed to obtain this system is $N/2$.

An analysis is also provided for the case $K = N-1$. In this case, changing the component for one part changes the performance contributions of each selected component. Kauffman argues analytically that when N tends toward infinity, the replacement process results in a locally-stable system whose expected performance approaches 1/2. It is also shown that the expected number of locally-stable systems is $2^N/(N+1)$ and that the expected number of replacements needed to reach a locally-stable system is $\ln(N - 1)$.

Kauffman then uses computer simulations to determine the average performance of a locally-stable system for different values of K. Those results indicate that, as N gets large, for positive values of K up to about 8, the expected performance exceeds the performance of 2/3 associated with $K = 0$. But then, as K increases, the expected performance of a locally-stable system decreases toward 1/2. Kauffman refers to this phenomenon—of decreasing performance associated with increasing interaction—as the *complexity catastrophe* (see Figure 2).

Intuitively, the complexity catastrophe is due to a trade-off that arises as K increases. The larger the value of K, the greater the number (2^{K+1}) of possible values for the performance contributions of each chosen component, thus resulting in more choices for components that could result in individual performance contributions close to 1. However, as K increases, there are more conflicts between the chosen components in

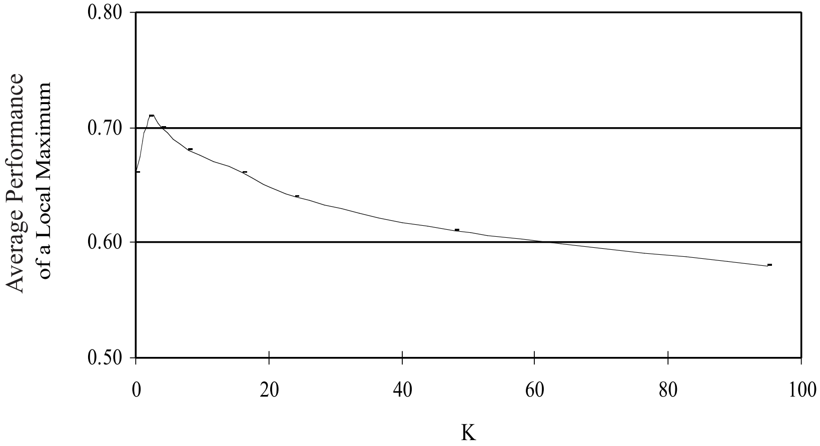

Figure 2. The complexity catastrophe in Kauffman's *NK* model ($N = 96$).

that whatever component is chosen for part i, that component is likely to benefit the performance contributions of some chosen components while being detrimental to the performance contributions of other components. The simulation results obtained from the *NK* model indicate that for small positive values of K, the benefits of having more choices for the performance contributions of the chosen components outweigh the few conflicts. However, as K increases, the negative effects of the increasing number of conflicts dominate the benefits of having more choices for the performance contributions of the chosen components, thus resulting in no more than average overall performance for each component and hence for the system as a whole.

A formal proof of the complexity catastrophe for the *NK* model is provided in [6]. However, several modifications of the basic model have demonstrated that the complexity catastrophe can be overcome (e.g., [7], in which the performance of a system is computed as a weighted average of the individual contributions). Subsequently, [8] shows that the complexity catastrophe is linked critically to a number of underlying mathematical assumptions inherent in the *NK* model. Those assumptions are identified explicitly and it is shown that, by relaxing any one of them, it is possible to create a slightly modified—and in some cases, a more realistic—model in which the complexity catastrophe is attenuated.

2.2 A general model of systems with complex interacting components

The following more general model is proposed now for studying systems with N parts. To that end, suppose that for each part i, it is possible to choose one of c_i alternative components, numbered 0 through $c_i - 1$, so, let

$$Z(c_i) = \{\text{integers } x\colon 0 \le x \le c_i - 1\}.$$

A system is then represented by an N-vector $\mathbf{x} = (x_1, \dots, x_N)$, in which each $x_i \in Z(c_i)$. Equivalently, letting $\mathbf{c} = (c_1, \dots, c_N)$,

$$\mathbf{x} \in Z(\mathbf{c}) = Z(c_1) \times \cdots \times Z(c_N).$$

To capture the interactions among the components in a system, for each part i, let K_i be the number of other parts that affect the performance contribution of part i and let S_i be the subscripts of those parts whose components affect the performance contribution of the component chosen for part i. In this paper, it is assumed that K_i and S_i do not depend on $\mathbf{x}$, although such an extension is possible and may be applicable for certain systems. Notationally,

S_i = {integers $j \neq i\colon 1 \le j \le N$ and the component chosen for part j affects the performance contribution of the component chosen for part i},
$K_i = |S_i|$,
$\mathbf{x}^i$ = the vector consisting of x_i and the K_i components of $\mathbf{x}$ indexed by the subscripts in S_i.

Turning to the performance measure, the interactions among the chosen components, represented by the values of S_i and K_i, are incorporated in the fact that p_i is a function of $\mathbf{x}^i$, that is,

$p_i(\mathbf{x}^i)$ = the contribution to performance of the component chosen for part i, which depends on the components of $\mathbf{x}$ in $\mathbf{x}^i$.

The performance of the system $\mathbf{x}$, denoted by $p(\mathbf{x})$, is then obtained by combining the contributions $p_i(\mathbf{x}^i)$ of the individual selected components through the use of a function g of N real numbers. That is,

$$p(\mathbf{x}) = g(p_1(\mathbf{x}^1), \dots, p_N(\mathbf{x}^N)). \tag{2}$$

Some typical examples of the function g include:

$$g(\mathbf{y}) = \frac{\sum_{i=1}^N y_i}{N}, \quad g(\mathbf{y}) = \max_{i=1,\dots,N} y_i, \quad g(\mathbf{y}) = \min_{i=1,\dots,N} y_i, \quad g(\mathbf{y}) = \left(\prod_{i=1}^N y_i\right)^{1/N}.$$

In summary, given values for N and the number of components c_i available for each part i, the proposed model consists of all N-vectors $\mathbf{x} = (x_1, \dots, x_N)$, in which each $x_i \in Z(c_i)$, together with the interaction data consisting of K_i, S_i, and the N functions $p_i(\mathbf{x}^i)$ together with the function g used to combine the $p_i(\mathbf{x}^i)$, as given in equation (2), to provide the overall performance $p(\mathbf{x})$ of the system $\mathbf{x}$. The objective is to determine values for the components $x_1, \ldots, x_N$ so that the resulting system $\mathbf{x} = (x_1, \ldots, x_N)$ has the best performance, according to equation (2).

Note that Kauffman's *NK* model is a special case of that proposed here in which (a) $c_i = 2$, for each $i = 1, \ldots, N$; (b) $K_i = K$, for each $i = 1, \ldots, N$; (c) $S_i = \{i - K/2, \ldots, i, \ldots, i + K/2\}$ if K is even and $S_i = \{i - \lfloor K/2 \rfloor - 1, \ldots, i, \ldots, i + \lfloor K/2 \rfloor\}$ if K is odd; and (d) $p(\mathbf{x}) = \sum_{i=1}^{N} p_i(\mathbf{x}^i)/N$.

3. Computational complexity results and heuristics

The computational complexity of finding the best system is addressed in this section. To that end, it is assumed that, for each $\mathbf{x} \in Z(\mathbf{c})$, each function $p_i(\mathbf{x}^i)$ as well as g return a rational number between 0 and 1 in polynomial time, in terms of N. The associated decision problem is now stated formally.

> *NK*: Given an integer $N > 0$ and, for each $i = 1, \ldots, N$, positive integers c_i, sets $S_i \subseteq \{1, \ldots, N\} - \{i\}$ with $K_i = |S_i|$, polynomially computable functions $p_i(\mathbf{x}^i)$ and g, and a rational number $C > 0$, does there exist an N-vector $\mathbf{x} \in Z(\mathbf{c})$ such that $g(p_1(\mathbf{x}^1), \ldots, p_N(\mathbf{x}^N)) \geq C$?

Theorems 3.1 and 3.2, whose proofs are given in appendix A, show that two special cases of the foregoing problem are strongly *NP*-complete, so the *NK* problem is also strongly *NP*-complete.

Theorem 3.1. For any positive integer Q, the following problem is strongly *NP*-complete:

> *NK*(Q): Given $N > Q$, $K = N - Q$, $c_i = 2$, $S_i = \{i + 1 - \lfloor (i+1)/(N+1) \rfloor N, \ldots, i + K - \lfloor (i+K)/(N+1) \rfloor N\}$, the functions $p_i(\mathbf{x}^i)$, for $i = 1, \ldots, N$, and a rational number $C > 0$, does there exist a binary N-vector $\mathbf{x}$ such that $\sum_{i=1}^{N} p_i(\mathbf{x}^i)/N \geq C$?

Theorem 3.2. Let α be a rational number with $0 < \alpha < 1$. The following problem is strongly *NP*-complete.

> *NK*(α): Given N, $K = \min\{\lceil \alpha N \rceil, N - 1\}$, $c_i = 2$, $S_i = \{i + 1 - \lfloor (i+1)/(N+1) \rfloor N, \ldots, i + K - \lfloor (i+K)/(N+1) \rfloor N\}$, the functions $p_i(\mathbf{x}^i)$, for $i = 1, \ldots, N$, and a rational number $C > 0$, does there exist a binary N-vector $\mathbf{x}$ such that $\sum_{i=1}^{N} p_i(\mathbf{x}^i)/N \geq C$?

Theorems 3.1 and 3.2 mean that the general *NK* problem is strongly *NP*-complete. Therefore, it is only possible to find the best system, by explicit enumeration, when N and $c_1, \ldots, c_N$ are relatively small. Another alternative is to develop polynomial algorithms for finding the best system for certain special cases, as is done in section 3.1. Then, in section 3.2, heuristics are proposed for the general problem.

3.1 Polynomial-solvable special cases

In this section, two special cases of the general problem are presented for which it is possible to develop polynomial algorithms to obtain the best system.

3.1.1 A polynomial algorithm when each $K_i = 0$

For the case in which each $K_i = 0$, that is, when each $S_i = \emptyset$, the performance contribution of each component i in the system $\mathbf{x}$ is determined only by component i, that is,

$$p_i(\mathbf{x}^i) = p_i(x_i).$$

In this case, the performance of the system $\mathbf{x}$ is

$$p(\mathbf{x}) = g(p_1(x_1), \ldots, p_N(x_N)).$$

Finding a system with the best overall performance normally involves considering an exponential number of possible systems. Definition 3.1 and Theorem 3.3 reduce this task to a polynomial amount of work by providing a condition under which a best system is found by identifying, for each part i, the component that maximizes the contribution $p_i(x_i)$.

Definition 3.1. A function $g : R^N \to R^1$ is *monotonically nondecreasing* if and only if whenever $\mathbf{a}, \mathbf{b} \in R^N$ with $\mathbf{a} \le \mathbf{b}$, it follows that $g(\mathbf{a}) \le g(\mathbf{b})$.

Theorem 3.3. If $g : R^N \to R^1$ is monotonically nondecreasing, then the system $\mathbf{x}^*$ with

$$p_i(x_i^*) = \max_{x_i \in Z(c_i)} p_i(x_i) \quad \text{for each } i = 1, \ldots, N$$

satisfies

$$\max_{\mathbf{x} \in Z(\mathbf{c})} g(p_1(x_1), \ldots, p_N(x_N)) = g(p_1(x_1^*), \ldots, p_N(x_N^*)). \tag{3}$$

Proof. Let $\bar{\mathbf{x}}$ be a system that achieves the best possible performance. By definition, then, the performance of $\bar{\mathbf{x}}$ is better than the performance of the system $\mathbf{x}^*$, that is,

$$g(p_1(\bar{x}_1), \ldots, p_N(\bar{x}_N)) \ge g(p_1(x_1^*), \ldots, p_N(x_N^*)). \tag{4}$$

For the other inequality, note that, by definition of $\mathbf{x}^*$,

$$p_i(\bar{x}_i) \le p_i(x_i^*), \quad \text{for each } i = 1, \ldots, N.$$

By the monotonicity of g, it follows that

$$g(p_1(\bar{x}_1), \ldots, p_N(\bar{x}_N)) \le g(p_1(x_1^*), \ldots, p_N(x_N^*)). \tag{5}$$

The desired conclusion follows by combining equations (4) and (5). ■

Theorem 3.3 applies to each of the following monotonically nondecreasing functions (where w_i are nonnegative real numbers that sum to 1 and, in the last two cases, g is defined on the nonnegative orthant of R^N):

$$\begin{aligned}
g(\mathbf{y}) &= \sum_{i=1}^{N} w_i y_i,\\
g(\mathbf{y}) &= \max_{i=1,\dots,N} y_i,\\
g(\mathbf{y}) &= \min_{i=1,\dots,N} y_i,\\
g(\mathbf{y}) &= \prod_{i=1}^{N} y_i,\\
g(\mathbf{y}) &= \left(\prod_{i=1}^{N} y_i\right)^{1/N}.
\end{aligned}$$

Furthermore, the amount of work needed to find the optimal system $\mathbf{x}^*$ in Theorem 3.3 depends on the amount of work needed to evaluate the individual contributions, $p_i(x_i)$. Specifically, if $q_i(N)$ is the polynomial amount of work needed to evaluate p_i and $d = \deg(\sum_{i=1}^{N} q_i(N))$, then the amount of work needed to find the optimal system $\mathbf{x}^*$ is $O(N^d)$.

3.1.2 A polynomial algorithm when g and S_i have special properties

For the special case in which $c_i = 2$ for all i, the contribution of the component chosen for part i depends on the components chosen in the K subsequent positions, wrapping around when necessary, and the system performance is taken to be the average of the individual contributions, in [4] it is shown that for a fixed value of K, the best system can be found in polynomial time by solving 2^K longest-path problems in an appropriate directed network. The approach from [4] is now generalized to measures of performance that combine the individual contributions in a specific way. For notational simplicity, it is still assumed that $c_i = 2$ for all i, however, all results developed in this section can be extended in a straightforward manner for general $\mathbf{c}$. To that end, let $B = \{0, 1\}$ and, for a vector $\mathbf{x} \in B^N = B \times \cdots \times B$, let $\mathbf{x}^i$ denote the vector that consists of x_i and the subsequent K components of $\mathbf{x}$, wrapping around when necessary, that is,

$$\mathbf{x}^i = (x_{w(i)}, \dots, x_{w(i+K)}), \quad \text{where} \quad w(j) = \begin{cases} j, & \text{if } 1 \le j \le N \\ j - N, & \text{if } j > N. \end{cases}$$

A dynamic programming approach is now proposed for solving the following problem

$$\begin{array}{ll} \text{Maximize} & g(p_1(\mathbf{x}^1), \dots, p_N(\mathbf{x}^N)) \\ \text{subject to} & \mathbf{x} \in B^N. \end{array}$$

Specifically, the approach is to fix the first K positions of $\mathbf{x}$ to, say, $\bar{\mathbf{x}} \in B^K$, and then solve the following problem $NK(\bar{\mathbf{x}})$ for the remaining components of $\mathbf{x}$:

$$\begin{array}{ll} \text{Maximize} & g(p_1(\mathbf{x}^1), \ldots, p_N(\mathbf{x}^N)) \\ \text{subject to} & \mathbf{x} \in B^N \text{ and } x_1 = \bar{x}_1, \ldots, x_K = \bar{x}_K \end{array} \qquad NK(\bar{\mathbf{x}}).$$

There are 2^K possible values for $\bar{\mathbf{x}}$, so, whichever of the 2^K problems $NK(\bar{\mathbf{x}})$ has the best objective function value provides the solution to the original NK problem. Because K is assumed to be independent of N, the complexity of this algorithm is that of the algorithm for solving $NK(\bar{\mathbf{x}})$. As seen in what follows, under reasonable assumptions, if $q_i(N)$ is the polynomial amount of work needed to evaluate p_i and $d = \deg(\sum_{i=1}^N q_i(N))$, then the complexity of the algorithm for solving $NK(\bar{\mathbf{x}})$ is $O(N^d)$.

To develop a dynamic programming recursion, certain decomposability properties are now proposed for the system performance function g. In a more general framework, sufficient conditions are developed in [9] for the optimality of a dynamic-programming recursion in problems where the objective function is replaced by appropriate preference relations among decisions. The desired property here is that g be computable by first combining a function α_N of $p_N(\mathbf{x}^N)$ with a function α_{N-1} of $p_{N-1}(\mathbf{x}^{N-1})$, then combining this result with a function α_{N-2} of $p_{N-2}(\mathbf{x}^{N-2})$ and so on, as described in Definition 3.2.

Definition 3.2. Given a nonempty subset $D \subseteq R^1$, a function

$$g : D \times \cdots \times D \subseteq R^N \to D$$

is *pairwise decomposable* if there are functions $h : D \times D \to D$ and $\alpha_i : D \to D$ for $i = 1, \ldots, N$ with h monotonically nondecreasing on $D \times D$ such that for all $\mathbf{y} \in D \times \cdots \times D$,

$$g(y_1, \ldots, y_N) = h(\alpha_1(y_1), \ldots, h(\alpha_{N-2}(y_{N-2}), h(\alpha_{N-1}(y_{N-1}), \alpha_N(y_N)))).$$

For example, using $D = \{y \in R^1 : y \geq 0\}$, each of the following functions is pairwise decomposable.

$g(y_1, \ldots, y_N)$	$\sum_{i=1}^{N} \frac{y_i}{N}$	$\sum_{i=1}^{N} w_i y_i$	$\max_{i=1,\ldots,N} y_i$	$\min_{i=1,\ldots,N} y_i$	$\prod_{i=1}^{N} y_i$	$\left(\prod_{i=1}^{N} y_i\right)^{1/N}$
$h(a,b)$	$a+b$	$a+b$	$\max\{a,b\}$	$\min\{a,b\}$	$a*b$	$a*b$
$\alpha_i(c)$	$\frac{c}{N}$	$w_i c$	c	c	c	$c^{1/N}$

The dynamic programming approach for solving each $NK(\bar{\mathbf{x}})$ involves N stages. For each stage $i = 1, \ldots, N$, the state of the problem is described by a vector $\hat{\mathbf{x}} \in B^K$, that is,

$\hat{\mathbf{x}} \in B^K$ = the state of the system in stage i ($i = 1, \ldots, N$), with $\hat{\mathbf{x}}$ representing the K values $x_{w(i)} = \hat{x}_1, \ldots, x_{w(i+K-1)} = \hat{x}_K$.

If g is pairwise decomposable with functions h and α_i, then the decision to be made at each stage i is the value of $x_{w(i+K)}$ (either 0 or 1) so as to maximize the following value function:

$V_i(\hat{\mathbf{x}})$ = the maximum, over all values of $x_{w(i+K)}, \ldots, x_{w(N+K)}$, of $h(\alpha_i(p_i(\mathbf{x}^i)), \ldots, h(\alpha_{N-1}(p_{N-1}(\mathbf{x}^{N-1})), \alpha_N(p_N(\mathbf{x}^N)))$, given that the values of $x_{w(i)} = \hat{x}_1, \ldots, x_{w(i+K-1)} = \hat{x}_K$.

When solving the problem $NK(\bar{\mathbf{x}})$, the value of $x_{w(N+i)} = x_i$ is fixed to $\bar{x}_i$ for $i = 1, \ldots, K$, so the effective state space and action sets must be modified to include these restrictions to ensure that the dynamic programming equations developed below are also valid as boundary conditions. Specifically, at any stage i for which $i + K > N$, the decision to be made applies to position $w(i + K) = i + K - N$, in which the value of x_{i+K-N} is fixed to $\bar{x}_{i+K-N}$. Therefore the only available action in this stage is to set $x_{w(i+K)} = \bar{x}_{i+K-N}$. In summary, let $A_i(\bar{\mathbf{x}})$ denote the set of available decisions in stage i of the problem $NK(\bar{\mathbf{x}})$. Then

$$A_i(\bar{\mathbf{x}}) = \begin{cases} \{0, 1\}, & \text{if } i + K \leq N \\ \{\bar{x}_{i+K-N}\}, & \text{if } i + K > N \end{cases}$$

Using analogous reasoning, the state space $S_i(\hat{\mathbf{x}})$ at stage i of the problem $NK(\bar{\mathbf{x}})$ is restricted so that any component either in the beginning or at the end (or both) of the state vector $\hat{\mathbf{x}}$ that coincides with x_j for some $j = 1, \ldots, K$ is fixed to the value $\bar{x}_j$. In summary, the following four cases must be considered in defining $S_i(\hat{\mathbf{x}})$.

1. $i > K, i \leq N - K + 1$. Then $S_i(\hat{\mathbf{x}}) = B^K$.
2. $i > K, i > N - K + 1$. Then $S_i(\hat{\mathbf{x}}) = \{\hat{\mathbf{x}} \in B^K : \hat{x}_j = \bar{x}_{i+j-1-N}, j = N - i + 2, \ldots, K\}$.
3. $i \leq K, i \leq N-K+1$. Then $S_i(\hat{\mathbf{x}}) = \{\hat{\mathbf{x}} \in B^K : \hat{x}_j = \bar{x}_{i+j-1}, j = 1, \ldots, K-i+1\}$.
4. $i \leq K, i > N-K+1$. Then $S_i(\hat{\mathbf{x}}) = \{\hat{\mathbf{x}} \in B^K : \hat{x}_j = \bar{x}_{i+j-1}, j = 1, \ldots, K-i+1, \hat{x}_j = \bar{x}_{i+j-1-N}, j = N - i + 2, \ldots, K\}$.

Note that, depending on whether $2K + 1 \leq N$ or not, the range of i corresponding to one or more of the foregoing cases may be null. However, the foregoing descriptions are stated in sufficient generality to accommodate all combinations of K and i values.

The following recursion is used to work backward until obtaining $V_1(\bar{\mathbf{x}})$:

$$\begin{aligned} V_i(\hat{\mathbf{x}}) = \max\{&h(\alpha_i(p_i(\hat{\mathbf{x}}, 0)), V_{i+1}(\hat{x}_2, \ldots, \hat{x}_K, 0)), \\ &h(\alpha_i(p_i(\hat{\mathbf{x}}, 1)), V_{i+1}(\hat{x}_2, \ldots, \hat{x}_K, 1))\} \end{aligned}$$

for $i = 1, \ldots, N-1$, with the following boundary condition at stage N:

$$V_N(\hat{\mathbf{x}}) = \alpha_N(p_N(\hat{\mathbf{x}}, \bar{x}_K)).$$

On obtaining $V_1(\bar{\mathbf{x}})$ by the above equations, a system is created in which $x_1 = \bar{x}_1, \ldots, x_K = \bar{x}_K$ and $x_{K+1}, \ldots, x_N$ are the optimal decisions made at stages $1, \ldots, N-K$. This is proved formally in Theorem 3.4.

Theorem 3.4. If g is pairwise decomposable with functions h and $\alpha_1, \ldots, \alpha_N$ then, for any $\bar{\mathbf{x}} \in B^K$, $V_1(\bar{\mathbf{x}})$ is the optimal value of the problem $NK(\bar{\mathbf{x}})$.

Proof. Let $\mathbf{z} \in B^N$ solve $NK(\bar{\mathbf{x}})$, so $z_1 = \bar{x}_1, \ldots, z_K = \bar{x}_K$. Also, let $\mathbf{y} \in B^N$ be the solution obtained from the dynamic programming recursion, so $y_1 = \bar{x}_1, \ldots, y_K = \bar{x}_K$ and $y_{K+1}, \ldots, y_N$ are the optimal decisions made at stages $1, \ldots, N-K$.

Then, because $\mathbf{z}$ is optimal for $NK(\bar{\mathbf{x}})$, it follows that

$$g(p_1(\mathbf{z}^1), \ldots, p_N(\mathbf{z}^N)) \geq g(p_1(\mathbf{y}^1), \ldots, p_N(\mathbf{y}^N)).$$

It remains to show that

$$g(p_1(\mathbf{y}^1), \ldots, p_N(\mathbf{y}^N)) \geq g(p_1(\mathbf{z}^1), \ldots, p_N(\mathbf{z}^N)).$$

Notationally, for $\mathbf{y}$, let $\hat{\mathbf{y}}^i = (y_{w(i)}, \ldots, y_{w(i+K-1)})$ and similarly for $\mathbf{z}$. Then, from the fact that g is pairwise decomposable and $\mathbf{y}$ is obtained from the dynamic programming recursion,

$$\begin{aligned} g(p_1(\mathbf{y}^1), \ldots, p_N(\mathbf{y}^N)) &= h(\alpha_1(p_1(\mathbf{y}^1)), \ldots, h(\alpha_{N-2}(p_{N-2}(\mathbf{y}^{N-2})), \\ &\qquad h(\alpha_{N-1}(p_{N-1}(\mathbf{y}^{N-1})), \alpha_N(p_N(\mathbf{y}^N)))) \\ &= h(\alpha_1(p_1(\mathbf{y}^1)), V_2(\hat{\mathbf{y}}^2)) \\ &\geq h(\alpha_1(p_1(\mathbf{z}^1)), V_2(\hat{\mathbf{z}}^2)). \end{aligned}$$

But since

$$(\alpha_1(p_1(\mathbf{z}^1)), V_2(\hat{\mathbf{z}}^2)) \geq (\alpha_1(p_1(\mathbf{z}^1)), h(\alpha_2(p_2(\mathbf{z}^2)), V_3(\hat{\mathbf{z}}^3))),$$

it follows by the monotonicity of h that

$$h(\alpha_1(p_1(\mathbf{z}^1)), V_2(\hat{\mathbf{z}}^2)) \geq h(\alpha_1(p_1(\mathbf{z}^1)), h(\alpha_2(p_2(\mathbf{z}^2)), V_3(\hat{\mathbf{z}}^3))).$$

Continuing in this manner yields

$$\begin{aligned} g(p_1(\mathbf{y}^1), \ldots, p_N(\mathbf{y}^N)) &\geq h(\alpha_1(p_1(\mathbf{z}^1)), \ldots, h(\alpha_{N-1}(p_{N-1}(\mathbf{z}^{N1})), \\ &\qquad \alpha_N(p_N(\mathbf{z}^N)))) \\ &= g(p_1(\mathbf{z}^1), \ldots, p_N(\mathbf{z}^N)). \end{aligned}$$

The proof is now complete. ■

3.2 Heuristics for building an effective system

From the results in section 3.1, it is possible to determine the best system when N and $c_1, \ldots, c_N$ are small—by explicitly checking all possible systems—or when g is monotonic and each $K_i = 0$ (see Theorem 3.3), or when g is pairwise decomposable and when each $K_i = K$ is small—by using the dynamic programming approach developed in section 3.1.2.

For the general problem, which is known to be *NP*-complete, heuristics are needed to find a good system with a reasonable amount of computational effort.

One such heuristic is to find a locally-stable system using the replacement process described in section 2.1.2. Another alternative proposed in [8] is to use a j-replacement process, in which up to j components in the current system $\mathbf{x}$ are changed simultaneously. A new heuristic, based on the dynamic programming approach presented in section 3.1.2, is to use the replacement process to obtain a locally-stable system $\mathbf{x}$ and then to solve a single dynamic programming problem of the type described in section 3.1.2, as follows.

Step 1. Use the replacement process to find a locally-stable system $\mathbf{x}$.

Step 2. Using $\mathbf{x}$, obtain a system $\mathbf{x}'$ by solving the dynamic programming problem in which the first K components are those of $\mathbf{x}$. If $\mathbf{x}' = \mathbf{x}$, stop. Otherwise, set $\mathbf{x} = \mathbf{x}'$ and repeat.

Unfortunately, the complexity of the replacement process in Step 1 is $O(2^N)$, however, such worst-case behavior is not experienced in the simulations described in section 4.3. As in the *NK* model, as K increases, the number of local maxima increases quickly and hence the number of replacements needed to obtain a local maximum decreases quickly.

These simulations are used to compare the average performance of a system obtained by the foregoing heuristic to that obtained by the single replacement process. Because the dimension of the state space is 2^K, the heuristic was run only for $K = 2, 4$, and 8. As seen from the results in Figure 3 for N = 24, 48, 96, and various values of K, the benefits from applying the dynamic-programming heuristic are greater for larger values of N and K.

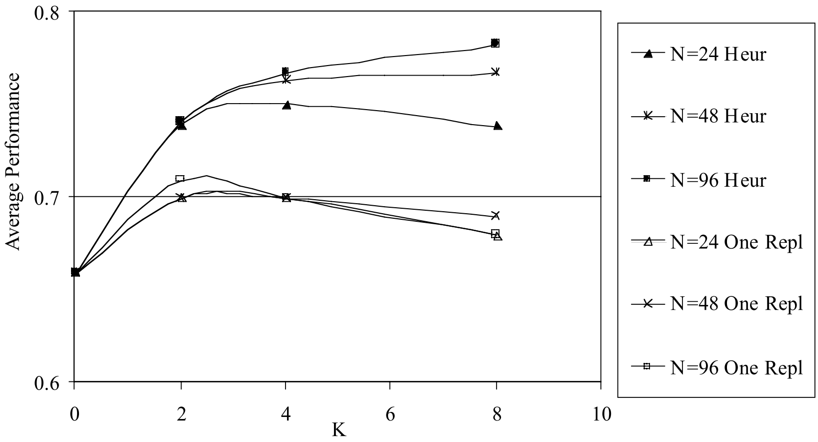

Figure 3. Average performance of a locally-stable system obtained by the heuristics in section 3.2.

4. Analytical and simulation results on the expected performance of the system obtained by the replacement process

In this section, the probabilistic behavior of the one-replacement process, described in section 2.1.2, is explored. The objective is to estimate the average performance of a locally-stable system when the performance contributions are random. Specifically, assume that the contribution $p_i(\mathbf{x})$ of the component chosen for part i to the overall performance of the system is a continuous random variable taking values between $-\infty$ and $+\infty$. In this case, a value greater than 0 indicates an above-average contribution to performance and a value less than 0 indicates a below-average contribution. Each such random variable is assumed to follow a probability density function $f_i(y)$ with finite mean μ_i and standard deviation σ_i. The corresponding cumulative distribution function is denoted by $F_i(y)$. It is also assumed that the values of $p_i(\mathbf{x}^i)$ are generated independently from the associated probability distributions.

Thinking now of the contributions to performance as random variables, it is of interest to find the expected value—and possibly higher moments—of the performance of the best system, that is, to find

$$M_N^{(m)} = E\left[\left(\max_{\mathbf{x}\in Z(\mathbf{c})} p(\mathbf{x})\right)^m\right]. \tag{6}$$

In particular, the variance of the performance can be derived from the second moment. For simplicity, the notation M_N is used when $m = 1$.

It should be emphasized here that the focus of this section is on the probabilistic analysis of the expected performance of a locally-stable system obtained by the replacement process described in section 2.1.2. Although related, this analysis differs from the usual approach in stochastic optimization, where a configuration (system) is sought that has the maximum expected performance when the performance contributions follow a specified probability distribution. The approach here is appropriate in situations where the entire landscape of performance contributions becomes known to the controller at the time of system optimization, whereas, the expected-value maximization is applicable in situations where decisions on system configuration are made before uncertainty is completely resolved.

It is also easy to see that the results of the previous sections can be recast in a stochastic optimization framework when the individual performance contributions $p_i(\mathbf{x})$ are random variables with mean $\mu_i(\mathbf{x})$, g has a weighted-sum form, and the system configuration $\mathbf{x}$ is to be determined before any information on the actual values of performance contributions is available. In this case, because of the linearity of the expectation operator, the expected performance of the system $\mathbf{x}$ is equal to $g(p_1(\mu_1(\mathbf{x})), \ldots, p_N(\mu(\mathbf{x}))$ and the analysis of section 3 is applicable. Of course, when individual component decisions are made sequentially and the uncertainty regarding performance is gradually resolved at the same time, more general stochastic optimization methods—such as stochastic programming or Markov decision control—are appropriate.

Analytical results are derived now for two special cases. Because in several cases below the results are asymptotic in nature as $N \to \infty$, the dependence of g on N is denoted explicitly to avoid confusion.

4.1 Analytical results when each $K_i = 0$

When each $K_i = 0$, the contribution of part i depends only on the component chosen for that part, that is, $p_i(\mathbf{x}^i) = p_i(x_i)$. In this case,

$$M_N^{(m)} = E\left[\left(\max_{\mathbf{x}\in Z(\mathbf{c})} g_N(p_1(x_1), \ldots, p_N(x_N))\right)^m\right].$$

Finding the foregoing expected value in general requires performing the following steps.

General steps for finding the moments of the best system performance

1. Find the density and cumulative distribution functions of a random variable $Y = g_N(V_1, \ldots, V_N)$, where each $V_i = p_i(x_i)$ is a random variable with density function f_i and cumulative distribution function F_i, as defined above.

2. Find the density and cumulative distribution function of a random variable Z representing the maximum of $c_1, \ldots, c_N$ random variables following the distribution of Y found in Step 1.

3. Use the density function found in Step 2 to compute the moments of the random variable Z.

When g is monotonically nondecreasing, the work needed to find the expected value of the performance of the best system is simplified. This is because, from Theorem 3.3, the best system is obtained by finding each component x_i^* that maximizes p_i. As a result,

$$M_N^{(m)} = E\left[\left(g_N(\max_{x_1\in Z(c_1)} p_1(x_1), \ldots, \max_{x_N\in Z(c_N)} p_N(x_N))\right)^m\right]. \tag{7}$$

Doing so now involves the following steps.

General steps for finding the moments of the best system performance when g is monotonically nondecreasing

1. Find the density and cumulative distribution functions of each random variable Z_i, where Z_i is the maximum of c_i random variables with density function f_i and cumulative distribution function F_i.

2. Find the density and cumulative distribution functions of a random variable $Z = g_N(Z_1, \ldots, Z_N)$.

3. Use the results in Step 2 to compute $M_N^{(m)} = E[Z^m]$.

Further simplifications that allow closed-form analysis arise when the contributions to performance are independent and identically distributed (i.i.d.) random variables following the uniform [0,1] distribution, and the function g_N takes several special forms. For simplicity, it is assumed that exactly two components are available for each part.

If g_N is linear in each $p_i(\mathbf{x})$, then, because $E[g_N(p(\mathbf{x}))] = g_N(E[p(\mathbf{x})])$, the expected value on the right-hand side of equation (7) can be computed as follows:

$$E\left[g_N\left(\max_{x_1 \in \{0,1\}} p_1(x_1), \ldots, \max_{x_N \in \{0,1\}} p_N(x_N)\right)\right] = g_N(E[Z_1], \ldots, E[Z_N]).$$

Therefore, if g_N is monotonically nondecreasing and linear in each $p_i(\mathbf{x})$, the expected value of a global maximum $\mathbf{x}^*$ can be computed by performing the following steps.

General steps for finding the expected value of the best system when g is monotonically nondecreasing and linear

1. Let $Z_i = \max\{p_i(0), p_i(1)\}$, where each $p_i(x_i) \sim f_i$.

2. Compute the expected value $E[Z_i]$.

3. Compute $M_N = g_N(E[Z_1], \ldots, E[Z_N])$, which is the expected value of a global maximum.

Based on the foregoing discussion, it is possible to obtain analytical results when g_N has certain special forms that are monotonically nondecreasing—such as the weighted average, maximum, minimum, product, and geometric average of its arguments. To illustrate the computation of $M_N^{(m)}$ for each of these functions, suppose that $p_i(x_i)$ are i.i.d. uniform [0,1] random variables. Thus, the distribution function of $Z_i = \max\{p_i(0), p_i(1)\}$ is $F_{Z_i}(s) = P[p_i(x_i) < s]^2 = s^2$ and the probability density is $f_{Z_i}(s) = 2s$. From this it follows that $E[Z_i^k] = \int_0^1 s^k f_{Z_i}(s) ds = 2/(2+k)$.

Analysis: $Y = g_N(p_1(x_1), \ldots, p_N(x_N)) = \sum_{i=1}^N w_i p_i(x_i)$, where $w_1, \ldots, w_N$ are nonnegative constants that sum to 1. In this case g_N is monotonically nondecreasing and linear, thus,

$$M_N = \sum_{i=1}^{N} w_i \left(\frac{2}{3}\right) = \frac{2}{3}.$$

In particular, the expected value of a global maximum for the original NK model, in which each $w_i = 1/N$, is 2/3.

The analytic expression for $M_N^{(m)}$ in this case is too complicated to provide any insights. However, the variance of Z is easy to obtain, since Z is the sum of independent random variables:

$$\text{Var}(Z) = \sum_{i=1}^{N} w_i^2 \text{Var}(Z_i),$$

and since $\text{Var}(Z_i) = E[Z_i^2] - (E[Z_i])^2 = 1/2 - 4/9 = 1/18$ for all i, it follows that

$$\text{Var}(Z) = \sum_{i=1}^{N} \frac{w_i^2}{18}.$$

The asymptotic behavior of Var(Z) as N becomes large depends on the choice of the weights $w_1, \ldots, w_N$. In the case where each $w_i = 1/N$, $\text{Var}(Z) = 1/(18N)$ which converges to zero as N becomes large.

Analysis: $Y = g_N(p_1(x_1), \ldots, p_N(x_N)) = \max_{i=1,\ldots,N} p_i(x_i)$. In this case, because $Z = \max\{Z_1, \ldots, Z_N\}$ and $Z_i = \max\{p_i(0), p_i(1)\}$, Z is distributed as the maximum of $2N$ i.i.d. uniform random variables, so,

$$F_Z(s) = P[p_i(\mathbf{x}) < s]^{2N} = s^{2N} \quad \text{and} \quad f_Z(s) = 2Ns^{2N-1}, \ s \in [0,1].$$

Therefore,

$$M_N^{(m)} = E[Z^m] = \int_0^1 s^m f_Z(s) ds = \int_0^1 2Ns^{2N+m-1} ds = \frac{2N}{2N+m}.$$

It now follows that $\lim_{N\to\infty} M_N^{(m)} = 1$ for any m, therefore

$$\lim_{N\to\infty} E[Z] = 1, \quad \text{and} \quad \lim_{N\to\infty} \text{Var}(Z) = 0.$$

Analysis: $Y = g_N(p_1(x_1), \ldots, p_N(x_N)) = \min_{i=1,\ldots,N} p_i(x_i)$. In this case,

$$\begin{aligned} F_Z(s) &= 1 - P[Z > s] = 1 - P[Z_i > s]^N \\ &= 1 - [1 - F_{Z_i}(s)]^N = 1 - (1 - s^2)^N \end{aligned}$$

and so

$$f_Z(s) = 2Ns(1 - s^2)^{N-1}.$$

It is now possible to compute

$$M_N^{(m)} = E[Z] = \int_0^1 s^m f_Z(s) ds = N \int_0^1 s^m (1 - s^2)^{N-1} 2s ds.$$

Using the variable transformation $z = s^2$, we obtain

$$M_N^{(m)} = N \int_0^1 z^{m/2}(1-z)^{N-1} dz = NB\left(\frac{m}{2}+1, N\right),$$

where $B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx$ is the Beta integral. In this case $\lim_{N\to\infty} M_N^{(m)} = 0$ for any $m \geq 1$, thus

$$\lim_{N\to\infty} E[Z] = 0, \quad \text{and} \quad \lim_{N\to\infty} \text{Var}(Z) = 0.$$

Analysis: $Y = g_N(p_1(x_1), \ldots, p_N(x_N)) = \prod_{i=1}^N p_i(x_i)$. In this case, the statistical independence of the Z_i implies that

$$M_N^{(m)} = E[Z^m] = \prod_{i=1}^N E[Z_i^m].$$

Since $f_{Z_i}(s) = 2s$, it follows that $E[Z_i^m] = 2/(m+2)$, for $i = 1, \ldots, N$ therefore,

$$M_N^{(m)} = \left(\frac{2}{m+2}\right)^N,$$

and $\lim_{N\to\infty} M_N^{(m)} = 0$ for any $m \geq 1$. In particular,

$$\lim_{N\to\infty} E[Z] = 0, \quad \text{and} \quad \lim_{N\to\infty} \text{Var}(Z) = 0.$$

Analysis: $Y = g_N(p_1(x_1), \ldots, p_N(x_N)) = \left(\prod_{i=1}^N p_i(x_i)\right)^{1/N}$. In this case, from $E[Z_i^m] = 2/(m+2)$ and the fact that the Z_i are independent, it follows that

$$M_N^{(m)} = E[Z^m] = \prod_{i=1}^N E[Z_i^{m/N}] = \left(\frac{2}{\frac{m}{N}+2}\right)^N = \left(1 - \frac{\frac{m}{2}}{\frac{m}{2}+N}\right)^N.$$

Thus, $\lim_{N\to\infty} M_N^{(m)} = e^{-m/2}$. In particular,

$$\lim_{N\to\infty} E[Z] = e^{-1/2}, \quad \text{and} \quad \lim_{N\to\infty} \text{Var}(Z) = 0.$$

The results of computing $M_N = E[Z]$ and the associated limiting values for each of the foregoing functions g are summarized in Table 1, where $B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx$ is the Beta integral. Note that in all cases the asymptotic variance of Z is zero.

4.2 Mathematical analysis when each $K_i = N - 1$

For the case in which each $K_i = N - 1$, the performance contribution of each component in the system $\mathbf{x}$ is affected by every other component.

$Y = g_N(V_1, \ldots, V_N)$	$\sum_{i=1}^{N} w_i V_i$	$\max_{i=1,\ldots,N} V_i$	$\min_{i=1,\ldots,N} V_i$	$\prod_{i=1}^{N} V_i$	$\left(\prod_{i=1}^{N} V_i\right)^{1/N}$
$M_N = E[Z]$	$\frac{2}{3}$	$\frac{2N}{2N+1}$	$NB\left(\frac{3}{2}, N\right)$	$\left(\frac{2}{3}\right)^N$	$\left(1 - \frac{1}{1+2N}\right)^N$
$\lim_{N\to\infty} M_N$	$\frac{2}{3}$	1	0	0	$e^{-1/2}$

Table 1. Limiting values of the expected performance of the best system using different performance measures ($K = 0$).

In this case, it is possible to derive asymptotic results for the expected performance of a locally-stable system obtained by the replacement process. This is because, when each $K_i = N - 1$, the performances of the N one-replacement neighbors of $\mathbf{x}$ are independent random variables. This fact is useful for computing an approximate value for the expected value of a locally-stable system when N is large. To do so, note that a locally-stable system $\mathbf{x}^*$ has the property that its performance $p(\mathbf{x}^*)$ is greater than or equal to that of its N one-replacement neighbors. Letting Z be a random variable representing the performance of a locally-stable system and $Y_1, \ldots, Y_{N+1}$ the performances of the locally-stable system and its N neighbors, it follows that

$$Z = \max\{Y_1, \ldots, Y_{N+1}\}. \tag{8}$$

To find the expected value of Z as N becomes large, it is necessary to follow these steps.

1. Find the density and cumulative distribution function of Y_i. That is, use the probability density functions f_i and distribution functions F_i of $p_i(\mathbf{x})$, to find the probability density function f_Y and the distribution function F_Y of the performance of a system $\mathbf{x}$, namely, $Y = g(p_1(\mathbf{x}), \ldots, p_N(\mathbf{x}))$.

2. Use the functions f_Y and F_Y from Step 1 to compute the probability density function f_Z and the distribution function F_Z of the random variable Z. Because, as stated in equation (8), Z is the maximum of $N+1$ independent random variables following the distribution of Y, it follows that

$$F_Z(t) = [F_Y(t)]^{N+1} \tag{9}$$

$$f_Z(t) = \frac{d}{dt}F_Z(t) = (N+1)[F_Y(t)]^N f_Y(t). \tag{10}$$

3. Use the function f_Z to compute the expected value of Z, as follows:

$$M_N = E[Z] = \int_{-\infty}^{\infty} t f_Z(t)dt = \int_{-\infty}^{\infty} tN[F_Y(t)]^{N-1} f_Y(t)dt.$$

4. Finally, compute $\lim_{N\to\infty} E[Z]$.

$Y = g_N(y_1, \ldots, y_N)$	$\frac{1}{N}\sum_{i=1}^{N} y_i$	$\max_{i=1,\ldots,N} y_i$	$\min_{i=1,\ldots,N} y_i$	$\prod_{i=1}^{N} y_i$	$\left(\prod_{i=1}^{N} y_i\right)^{1/N}$
$\lim_{N\to\infty} M_N$	$\frac{1}{2}$	1	0	0	e^{-1}

Table 2. Limiting values of the expected performance of a locally-stable system using different performance measures ($K = N - 1$).

Although in principle higher moments of Z can also be expressed in terms of the foregoing steps, we have not been able to obtain analytical expressions as was done for the case where $K_i = 0$, because the analysis here is substantially more complicated.

It is important to note that the foregoing analysis is based on a locally-stable system chosen at random rather than a locally-stable system found by the replacement process. The latter is different from the former because not all locally-stable systems are equally likely to be found by the replacement process. This is due to the fact that a locally-stable system found by the replacement process depends on the starting system. In particular, some locally-stable systems may have large basins of attraction (i.e., many initial systems will result in that locally-stable system) while other locally-stable systems may have small basins of attraction. The inability to handle these different types of locally-stable systems analytically is noted clearly in [10]. Nevertheless, analytical results for the limit of M_N as $N \to \infty$ for the special forms of the function g presented in section 3.1.2 are summarized in Table 2 and derived in appendix B.

4.3 Computer simulations

Computer simulations are used now to verify the analytical results for the expected value of the performance of the best system when $K = 0$ presented in Table 1 in section 4.1 and for a locally-stable system when $K = N - 1$ presented in Table 2 in section 4.2. Simulations are also used to estimate the performance of a locally-stable system obtained by the replacement process for intermediate values of K, for which no analytical results are available. The results of the average performance of a locally-stable system obtained from 1000 randomly-generated problems—in which the performance contributions $p_i(\mathrm{x}^i)$ are drawn i.i.d. from the uniform [0, 1] distribution—are presented in Figures 4 through 7 for the cases where the system performance is the maximum, minimum, product, and geometric mean of the individual contributions, respectively. In all cases, the average performance of a locally-stable system is presented as a function of K, for $N = 24$, 48, and 96, except for the product case, where the graphs for $N = 8$ and 16 are presented as well.

These results are in agreement with the theoretical expected values for the extreme cases of $K = 0$ and $N - 1$ in Tables 1 and 2. Furthermore,

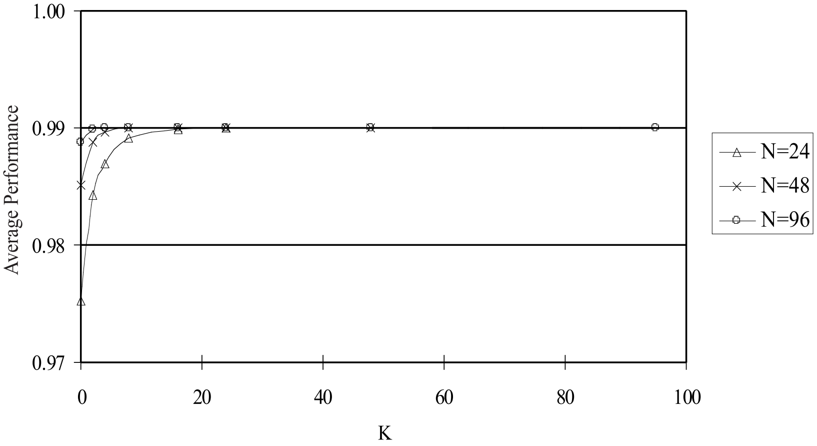

Figure 4. Average performance of a locally-stable system when system performance is the maximum of component performances.

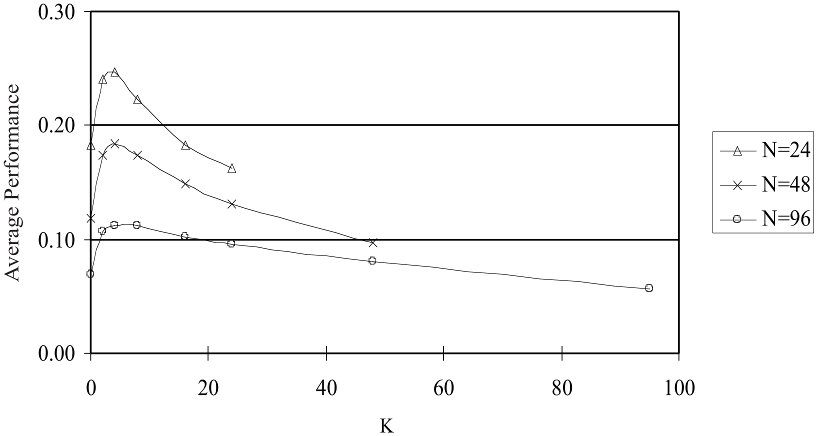

Figure 5. Average performance of a locally-stable system when system performance is the minimum of component performances.

in all cases except when the system performance is the maximum of the individual contributions, the complexity catastrophe arises, but to varying degrees. For example, the complexity catastrophe in Figure 7, corresponding to the geometric mean of the individual contributions, is virtually the same as that in the original *NK* model. This is because,

$$\text{when } p(\mathbf{x}) = \left(\prod_{i=1}^{N} p_i(\mathbf{x})\right)^{1/N}, \text{ it follows that } \ln(p(\mathbf{x})) = \sum_{i=1}^{N} \frac{p_i(\mathbf{x})}{N},$$

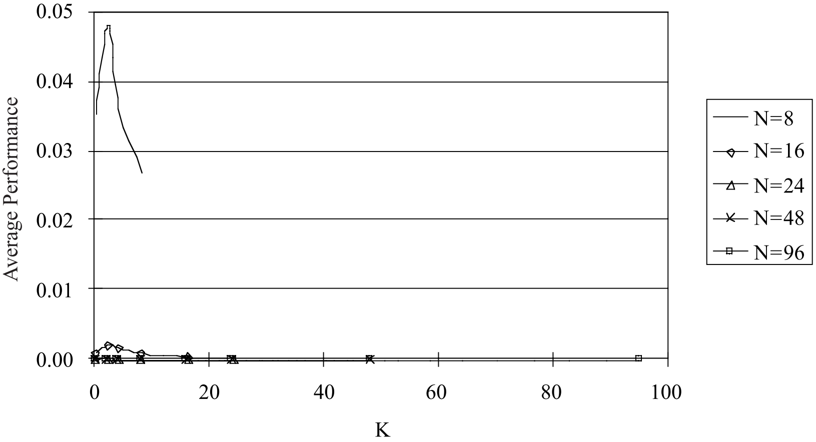

Figure 6. Average performance of a locally-stable system when system performance is the product of component performances.

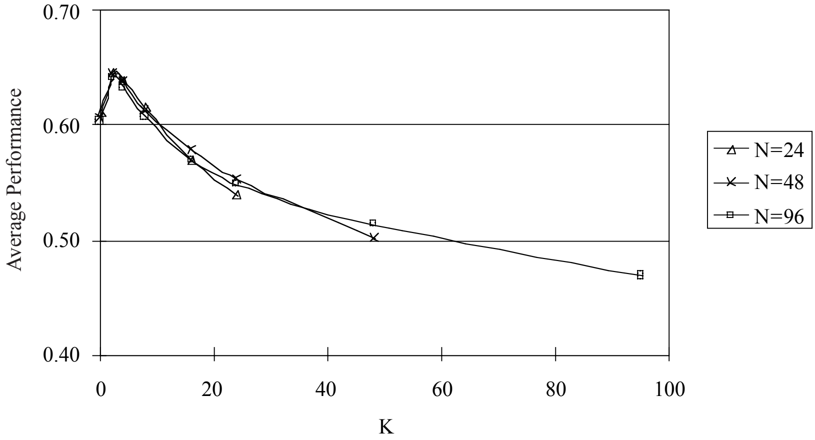

Figure 7. Average performance of a locally-stable system when performance is the geometric mean of component performances.

which is how system performance is computed in the original *NK* model. In contrast, in Figure 6—where system performance is computed as the product of the individual contributions—the complexity catastrophe is evident for small values of *N* (*N* = 8 and 16) but becomes unnoticable for large values of *N* because, in fact, system performance in this case quickly approaches 0, regardless of the value of *K*. In Figure 5—where system performance is based on the "weakest link" (i.e., the minimum of the individual contributions)—systems with fewer components benefit more from small amounts of interactions among the components than do larger systems. Finally, the fact that the complexity catastrophe is not evident in Figure 4—where system performance is computed as the maximum of the individual contributions—indicates that there are systems whose performance, in fact, benefits from increasing amounts of interactions among the components. This example supports the results in [8].

Conclusions and future research

A general model is proposed for obtaining insights into systems consisting of a number of parts that interact with each other in complex ways. To capture many aspects of real-world systems, the model allows one to choose, (a) for each part, one of a number of alternative components so that the resulting system has the best performance, according to some specific measure, (b) varying amounts of interaction among the components and which components affect other components, (c) the way in which the overall system performance is obtained from the individual performance contributions of the chosen components. The general problem of finding the best system is shown to be *NP*-complete and several polynomially-solvable special cases and heuristics are described.

When the contribution of each chosen component i depends on exactly K other chosen components $(0 \leq K \leq N - 1)$, the values of N and K provide a measure of the complexity of the system. Assuming that the individual performance contributions are random variables that follow a known distribution, conditions on how the performance of a system is computed are given under which it is possible to derive analytically the expected performance of a locally-stable system obtained by the replacement process. In particular, such results are given for $K = 0$, corresponding to a system in which the contribution to performance of each chosen component depends only on that component, and also for $K = N - 1$, corresponding to a system in which the contribution to performance of each chosen component depends on every other chosen component.

Computer simulations support the analytical results and also provide average performance values of locally-stable systems under different measures of overall performance. Some of these systems exhibit the property that for large values of N, the average performance of a locally-stable system associated with small positive values of K exceeds the performance associated with a locally-stable system obtained when $K = 0$. However, as the value of K, and hence the complexity of the system, increases, the average performance decreases to values below that associated with $K = 0$. It is important to note that this *complexity catastrophe*, as coined by Kauffman, is based on the collective mathematical assumptions underlying these specific models.

On the basis of these results, current research efforts are devoted to seeking ways to differentiate the individual parts to make these models more realistic. For instance, in some systems, certain parts contribute indirectly, rather than directly, to system performance. In the human genome, for example, regulatory genes contribute indirectly by controlling the behavior of other genes that do contribute directly to the fitness of the genome. As another example, in a team in a business organization, the leader contributes indirectly to the performance of the team by providing direction, setting goals, providing resources, and eliciting direct contributions from the individual team members. An area of current research is to build models that allow for such indirect contributions to see what impact they have on system performance.

Appendix

A. *NP*-completeness proofs for section 3

Proof of Theorem 3.1. Note first that the problem $NK(Q)$ is in NP because, given a binary N-vector $\mathbf{x}$, it is possible to check whether $\sum_{i=1}^{N} p_i(\mathbf{x}^i)/N \geq C$ in polynomial time.

To complete the proof, a reduction is given from the following problem of finding a largest independent set of nodes in a graph:

IS: Given a graph $G = (V, E)$ and an integer $B > 0$, does there exist a set S of nodes of G, no two of which are adjacent, such that $|S| \geq B$?

To that end, consider an instance of *IS* consisting of a graph $G = (V, E)$ and an integer $B > 0$. Let $M = |V|$ and define, for each binary M-vector $\mathbf{y}$, the following subset of nodes of $V = \{v_1, \ldots, v_M\}$:

$$S(\mathbf{y}) = \{v_i \in V : y_i = 1\}.$$

For a given positive integer Q, it is now possible to construct the associated instance of $NK(Q)$:

$$N = |V| + Q = M + Q, \quad K = N - Q = M, \quad C = \frac{B + |V|}{NP} = \frac{B + M}{NM}.$$

In this case, the contribution to performance p_i of each component i depends on $\mathbf{x}^i = (x_{w(i)}, \ldots, x_{w(i+K-1)})$. Then, the functions $p_1, \ldots, p_{Q+M}$ are defined as follows:

$$p_i(\mathbf{x}^i) = 0, \quad \text{for } i = 1, \ldots, Q - 1$$

$$p_Q(\mathbf{x}^Q) = \begin{cases} 1, & \text{if } S(x_{Q+1}, \ldots x_{Q+M}) \text{ is an independent set,} \\ 0, & \text{otherwise} \end{cases}$$

$$p_i(\mathbf{x}^i) = \frac{x_i}{M}, \quad \text{for } i = Q + 1, \ldots, Q + M.$$

Now suppose that $\mathbf{x}$ is a binary N-vector such that

$$\frac{\sum_{i=1}^{P+Q} p_i(\mathbf{x}^i)}{N} \geq C = \frac{B + |V|}{NM} = \frac{B + M}{NM}.$$

Equivalently,

$$\sum_{i=1}^{P+Q} p_i(\mathbf{x}^i) \geq \frac{B+M}{M}. \tag{A.1}$$

A corresponding solution to the *IS* problem is

$$S = S(x_{Q+1}, \ldots, x_{Q+M}).$$

To show that S is a solution to the *IS* problem, it is necessary to show that the nodes of S form an independent set and that $|S| \geq B$. Both of these follow from the fact that $\mathbf{x}$ satisfies equation (A.1). Specifically, because $p_i(\mathbf{x}^i) = 0$, for $i = 1, \ldots, Q-1$, equation (A.1) becomes

$$\sum_{i=1}^{M+Q} p_i(\mathbf{x}^i) = \sum_{i=Q}^{M+Q} p_i(\mathbf{x}^i) = p_Q(\mathbf{x}^Q) + \sum_{i=Q+1}^{M+Q} p_i(\mathbf{x}^i) \geq \frac{B+M}{M} = \frac{B}{M} + 1. \tag{A.2}$$

The only way for the sum in equation (A.2) to exceed $B/M + 1$ is for $p_Q(\mathbf{x}^Q)$ to be equal to 1 because, by construction, each $p_i(\mathbf{x}^i)$, for $i = Q+1, \ldots, Q+M$, can contribute at most $1/M$ to the sum. Hence, $p_Q(\mathbf{x}^Q) = 1$ which, by definition, means that S is an independent set of nodes. Furthermore, equation (A.2) now yields that S has at least B nodes because

$$|S| = \sum_{i=Q+1}^{Q+M} x_i = M \sum_{i=Q+1}^{Q+M} p_i(\mathbf{x}^i) \geq B.$$

Conversely, suppose that S is an independent set of nodes in G with $|S| \geq B$. Define the binary N-vector $\mathbf{x}$ as follows:

$$x_i = \begin{cases} 1, & \text{if } Q+1 \leq i \leq Q+M \text{ and } v_{i-Q} \in S \\ 0, & \text{otherwise.} \end{cases}$$

Then $p_Q(\mathbf{x}^Q) = 1$ and so

$$\begin{aligned} \sum_{i=1}^{Q+M} p_i(\mathbf{x}^i) &= \sum_{i=Q}^{Q+M} p_i(\mathbf{x}^i) = p_Q(\mathbf{x}^Q) + \sum_{i=Q+1}^{Q+M} p_i(\mathbf{x}^i) \\ &= 1 + \sum_{i=Q+1}^{Q+M} \frac{x_i}{M} = 1 + \frac{|S|}{M} \geq 1 + \frac{B}{M} = CN. \end{aligned}$$

The desired result follows on dividing the foregoing inequality through by N. ■

Proof of Theorem 3.2. Note first that this *NK* problem is in *NP* because, given a binary N-vector $\mathbf{x} \in B^N$, it is possible to check whether

$\sum_{i=1}^{N} p_i(x_i)/N \geq C$ in polynomial time, since evaluating each p_i is polynomial.

To complete the proof, a reduction from the problem of finding an independent set (*IS*) is given. So, let $G = (V, E)$ be a given graph with $M = |V|$ and let $B > 0$ be a given integer. These data are now used to construct the following instance of the *NK* problem:

$$N = \max\left\{\left\lceil \frac{M}{\alpha} \right\rceil, M+1\right\}, \quad K = \min\{\lceil \alpha N \rceil, N-1\}, \quad C = \frac{M+B}{NM}.$$

The functions $p_1, \ldots, p_N$ are now defined as follows:

$$\begin{aligned} p_i(\mathbf{x}^i) &= 0, \quad \text{for } i = 1, \ldots, N-M-1 \\ p_{N-M}(\mathbf{x}^{N-M}) &= \begin{cases} 1, & \text{if } S(x_{N-M+1}, \ldots, x_N) \text{ is an independent set,} \\ 0, & \text{otherwise} \end{cases} \\ p_i(\mathbf{x}^i) &= \frac{x_i}{M}, \quad \text{for } i = N-M+1, \ldots, N. \end{aligned}$$

Now suppose that $\mathbf{x}$ is a binary N-vector such that

$$\frac{\sum_{i=1}^{N} p_i(\mathbf{x}^i)}{N} \geq C = \frac{B+M}{NM}.$$

Equivalently,

$$\sum_{i=1}^{N} p_i(\mathbf{x}^i) \geq 1 + \frac{B}{M}. \tag{A.3}$$

A corresponding solution to the *IS* problem is

$$S = S(x_{N-M+1}, \ldots, x_N).$$

To show that S is a solution to *IS*, it is necessary to show that the nodes of S form an independent set and that $|S| \geq B$. Both of these follow from the fact that $\mathbf{x}$ satisfies equation (A.3). Specifically, because $p_i(\mathbf{x}^i) = 0$, for $i = 1, \ldots, N-M-1$, equation (A.3) becomes

$$\sum_{i=1}^{N} p_i(\mathbf{x}^i) = \sum_{i=N-M}^{N} p_i(\mathbf{x}^i) = p_{N-M}(\mathbf{x}^{N-M}) + \sum_{i=N-M+1}^{N} p_i(\mathbf{x}^i) \geq 1 + \frac{B}{M}. \tag{A.4}$$

The only way for the sum in equation (A.4) to exceed $B/M + 1$ is for $p_{N-M}(\mathbf{x}^{N-M})$ to be equal to 1 because, by construction, each $p_i(\mathbf{x}^i)$, for $i = N-M+1, \ldots, N$, can contribute at most $1/M$ to the sum. Hence, $p_{N-M}(\mathbf{x}^{N-M}) = 1$ which, by definition, means that S is an independent set of nodes. Furthermore, equation (A.4) now yields that S has at least B nodes because

$$|S| = \sum_{i=N-M+1}^{N} x_i = M \sum_{i=N-M+1}^{N} p_i(\mathbf{x}^i) \geq B.$$

Conversely, suppose that S is an independent set of nodes in G with $|S| \geq B$. Define the binary N-vector $\mathbf{x}$ as follows:

$$x_i = \begin{cases} 1, & \text{if } N - M + 1 \leq i \leq N \text{ and } v_{i-N+M} \in S \\ 0, & \text{otherwise.} \end{cases}$$

Then $p_{N-M}(\mathbf{x}^{N-M}) = 1$ and so

$$\begin{aligned} \sum_{i=1}^{N} p_i(\mathbf{x}^i) &= \sum_{i=N-M}^{N} p_i(\mathbf{x}^i) = p_{N-M}(\mathbf{x}^{N-M}) + \sum_{i=N-M+1}^{N} p_i(\mathbf{x}^i) \\ &= 1 + \sum_{i=Q+1}^{Q+M} \frac{x_i}{M} = 1 + \frac{|S|}{M} \geq 1 + \frac{B}{M} = CN. \end{aligned}$$

The desired result follows on dividing the foregoing inequality through by N, completing the proof. ■

B. Limiting results for the expected performance of the systems in section 4.2

Throughout this appendix, $K = N-1$, the individual contributions $p_i(\mathbf{x})$ are assumed to be i.i.d. uniform [0,1] random variables, Y is a random variable representing the performance of a random system, and Z is a random variable representing the performance of a random locally-stable system.

B.1 Analysis when system performance is the average of the individual contributions

For the case where the performance of a system is

$$p(\mathbf{x}) = \frac{\sum_{i=1}^{N} p_i(\mathbf{x})}{N},$$

the distribution of the performance of a random system for large N approaches the normal, by the central limit theorem. However, using this approximation must be combined with taking the limit of $E[Z]$ as $N \to \infty$, which entails the maximum of the N random variables $Y_1, \ldots, Y_N$. It is shown formally in [6] that $\lim_{N\to\infty} M_N = 1/2$, which corresponds to the complexity catastrophe discussed in section 2.1.2.

B.2 Analysis when system performance is the maximum of the individual contributions

When

$$p(\mathbf{x}) = \max_{i=1,\ldots,N} p_i(\mathbf{x}),$$

the distribution of the performance $Y = p(\mathbf{x})$ of a system $\mathbf{x}$ is $F_Y(s) = P[p_i(\mathbf{x}) < s]^N = s^N$, $s \in [0, 1]$, and the density function $f_Y(s) = Ns^{N-1}$. Therefore, the distribution and density functions of Z are

$$F_Z(s) = [F_Y(s)]^N = s^{N^2}$$

and

$$f_Z(s) = N^2 s^{N^2-1}.$$

It follows that

$$M_N = E[Z] = \int_0^1 s f_Z(s) ds = \int_0^1 N^2 s^{N^2} ds = \frac{N^2}{N^2+1},$$

and so $\lim_{N\to\infty} M_N = 1$.

B.3 Analysis when system performance is the minimum of the individual contributions

When

$$p(\mathbf{x}) = \min_{i=1,\ldots,N} p_i(\mathbf{x}),$$

the distribution of the performance $Y = p(\mathbf{x})$ of a system $\mathbf{x}$ is $F_Y(s) = 1 - P[Y > s] = 1 - P[p_i(\mathbf{x}) > s]^N = 1 - (1-s)^N, s \in [0, 1]$, and $f_Y(s) = N(1-s)^{N-1}$. Therefore, the distribution and density functions of Z are

$$F_Z(s) = F_Y(s)^N = [1 - (1-s)^N]^N$$

and

$$f_Z(s) = N^2[1 - (1-s)^N]^{N-1}(1-s)^{N-1}.$$

It follows that

$$M_N = E[Z] = \int_0^1 s f_Z(s) ds = \int_0^1 N^2 s[1 - (1-s)^N]^{N-1}(1-s)^{N-1} ds.$$

Using the transformation $y = (1-s)^N$, the foregoing integral becomes

$$\begin{aligned} M_N &= N \int_0^1 (1 - y^{1/N})(1-y)^{N-1} dy \\ &= N\left(\int_0^1 (1-y)^{N-1} dy - \int_0^1 y^{1/N}(1-y)^{N-1} dy\right) \\ &= 1 - NB\left(\frac{1}{N} + 1, N\right), \end{aligned}$$

where B denotes the Beta integral. It then follows that $\lim_{N\to\infty} M_N = 0$.

B.4 Analysis when system performance is the product of the individual contributions

When

$$Y = p(\mathbf{x}) = \prod_{i=1}^{N} p_i(\mathbf{x}),$$

it can be shown that $W \equiv -\ln Y$ follows the Gamma distribution with parameters N and 1. To see this, note that $W = \sum_{i=1}^{N} W_i$, where $W_i = -\ln(p_i(\mathbf{x}))$, and $p_i(\mathbf{x})$ follows the uniform [0,1] distribution. Therefore, the W_i are i.i.d. random variables with distribution function

$$F_{W_i}(w) = P[-\ln p_i(\mathbf{x}) \le w] = P[p_i(\mathbf{x}) \ge e^{-w}] = 1 - e^{-w}, \quad w \in [0, \infty].$$

That is, the W_i are exponentially distributed with rate 1. Thus, W, which is the sum of N i.i.d. exponential random variables, follows the Gamma distribution with parameters N and 1, so,

$$f_W(w) = \frac{w^{N-1} e^{-w}}{\Gamma(N)}, \quad w \ge 0, \tag{B.1}$$

where $\Gamma(N) = \int_0^1 w^{N-1} e^{-w} dw$ denotes the Gamma integral. Furthermore, using the relationship between the Gamma and Poisson distributions, the distribution function of W is

$$F_W(w) = \sum_{j=N}^{\infty} e^{-w} \frac{w^j}{j!}. \tag{B.2}$$

Having obtained the distribution of W, it is possible to derive the distribution of $Y = e^{-W}$ using standard properties for the distribution of a function of a random variable. Specifically, since $Y = e^{-W}$ is a monotonic function of W, it follows from [11] that

$$f_Y(y) = \frac{f_W(-\ln y)}{y} = \frac{(-\ln y)^{N-1} e^{\ln y}}{y \Gamma(N)} = \frac{(-\ln y)^{N-1}}{(N-1)!}.$$

Furthermore, using equation (B.2), the distribution function of Y is

$$\begin{aligned} F_Y(y) = P[Y \le y] &= P[W \ge -\ln y] = 1 - F_W(-\ln y) \\ &= \sum_{j=0}^{N-1} \frac{y(-\ln(y))^j}{j!}, \quad y \in [0,1]. \end{aligned}$$

To calculate $E[Z]$, where Z is the performance of a randomly selected locally-stable system, use equation (10) to obtain

$$\begin{aligned} M_N = E[Z] = \int_0^1 s f_Z(s) ds &= (N+1) \int_0^1 s f_Y(s) (F_Y(s))^N ds \\ &= \int_0^1 f_W(-\ln s)(1 - F_W(-\ln s))^N ds. \end{aligned}$$

Using the transformation $w = -\ln s$ as well as integration by parts, the above expression becomes

$$\begin{aligned}
M_N &= (N+1)\int_0^\infty e^{-w} f_W(w)(1-F_W(w))^N dw \\
&= \int_0^\infty e^{-w} d(1-F_W(w))^{N+1} \\
&= 1 - \int_0^\infty e^{-w}(1-F_W(w))^{N+1} dw \\
&\leq 1 - \int_0^\infty e^{-w}(1-(N+1)F_W(w))dw \\
&= (N+1)\int_0^\infty e^{-w} F_W(w)dw \\
&= (N+1)\int_0^\infty e^{-w} \sum_{j=N}^{\infty} \frac{e^{-w}w^j}{j!} dw \\
&= (N+1)\sum_{j=N}^{\infty} \frac{1}{j!}\int_0^\infty e^{-2w} w^j dw \\
&= (N+1)\sum_{j=N}^{\infty} \frac{1}{j!}\frac{j!}{2^{j+1}} \\
&= 2(N+1)\left(\frac{1}{2}\right)^{N+1}.
\end{aligned}$$

In summary,

$$0 \leq M_N \leq 2(N+1)\left(\frac{1}{2}\right)^{N+1}.$$

Taking the limit yields $\lim_{N\to\infty} M_N = 0$.

B.5 Analysis when system performance is the geometric mean of the individual contributions

When

$$Y = p(\mathbf{x}) = \left(\prod_{i=1}^{N} p_i(\mathbf{x})\right)^{1/N},$$

let $V \equiv -\ln Y = W/N$, where $W = \sum_{i=1}^{N} W_i$ and $W_i = -\ln(p_i(\mathbf{x}))$, as defined in the product case. From equations (B.1) and (B.2), the probability density and distribution functions of V are

$$f_V(v) = N f_W(Nv) = \frac{N(Nv)^{N-1}e^{-Nv}}{\Gamma(N)}, \quad v \geq 0,$$

and

$$F_V(v) = F_W(Nv) = \sum_{j=N}^{\infty} e^{-Nv}\frac{(Nv)^j}{j!},$$

from which it can be seen that V is a Gamma(N, N) random variable.

The distribution of Y is now given by

$$f_Y(y) = \frac{f_V(-\ln y)}{y} = \frac{N(-N\ln y)^{N-1}e^{N\ln y}}{y\Gamma(N)} = \frac{N(-N\ln y)^{N-1}y^{N-1}}{(N-1)!}$$

and

$$F_Y(y) = 1 - F_V(-\ln y) = \sum_{j=0}^{N-1}\frac{(-N\ln y)^j y^N}{j!} \quad y \in [0,1]. \tag{B.3}$$

Furthermore,

$$E[Y] = E[e^{-V}] = \int_0^{\infty} e^{-v}f_V(v)dv = \left(1+\frac{1}{N}\right)^{-N}.$$

Asymptotic lower and upper bounds are derived now for $M_N = E[Z]$, where $Z = \max\{Y_1, \ldots, Y_{N+1}\}$ is the performance of a locally-stable system and $Y_1, \ldots, Y_{N+1}$ are i.i.d. random variables with the distribution of Y defined by equation (B.3).

Observe first that

$$E[Z] = E[\max\{Y_1, \ldots, Y_{N+1}\}] \geq \max\{E[Y_1], \ldots, E[Y_{N+1}]\} = \left(1+\frac{1}{N}\right)^{-N},$$

therefore,

$$\liminf_{N\to\infty} M_N \geq e^{-1}. \tag{B.4}$$

It is next shown that $\limsup_{N\to\infty} M_N \leq e^{-1}$. The proof makes use of large-deviation type inequalities [12] but the result is independently derived here for completeness. Because $0 \leq Z \leq 1$, $M_N = E[Z] = \int_0^1 P[Z \geq u]du$ and, using equation (9),

$$M_N = \int_0^1 (1 - F_Z(u))du = 1 - \int_0^1 F_Z(u)du = 1 - \int_0^1 (F_Y(u))^{N+1}du.$$

Choose ϵ with $0 < \epsilon < e - 1$, so, $0 < e^{-1}(1+\epsilon) < 1$. Then, since $F_Y(u)$ is nondecreasing and nonnegative,

$$\begin{aligned} M_N &\leq 1 - \int_{e^{-1}(1+\epsilon)}^{1} (F_Y(u))^{N+1}du \\ &\leq 1 - [F_Y(e^{-1}(1+\epsilon))]^{N+1}(1 - e^{-1}(1+\epsilon)) \\ &= 1 - [1 - \bar{F}_Y(e^{-1}(1+\epsilon))]^{N+1}(1 - e^{-1}(1+\epsilon)) \\ &\leq 1 - [1 - (N+1)\bar{F}_Y(e^{-1}(1+\epsilon))](1 - e^{-1}(1+\epsilon)), \end{aligned} \tag{B.5}$$

where $\bar{F}_Y(y) = 1 - F_Y(y)$.

Let $\theta = \ln(1+\epsilon)$. It then follows that, for any $t \geq 0$,

$$\begin{aligned}\bar{F}_Y(e^{-1}(1+\epsilon)) &= P[Y \geq e^{-(1-\theta)}] \\ &= P[\ln Y \geq -(1-\theta)] \\ &= P\left[-\sum_{i=1}^{N} W_i \geq -N(1-\theta)\right] \\ &= P\left[e^{-t\sum_{i=1}^{N} W_i} \geq e^{-Nt(1-\theta)}\right], \end{aligned} \tag{B.6}$$

where the W_i are i.i.d. exponentially distributed random variables with rate 1. The Markov inequality then yields that

$$\bar{F}_Y(e^{-1}(1+\epsilon)) \leq E[e^{-t\sum_{i=1}^{N} W_i}]e^{Nt(1-\theta)} = (E[e^{-tW_1}])^N e^{Nt(1-\theta)}.$$

Since W_1 follows an exponential distribution with rate 1,

$$E[e^{-tW_1}] = \int_0^\infty e^{-tw}e^{-w}dw = \frac{1}{t+1},$$

thus,

$$\bar{F}_Y(e^{-1}(1+\epsilon)) \leq e^{-Nh(t,\theta)}, \tag{B.7}$$

where $h(t,\theta) = \ln(t+1) - t(1-\theta)$. Since inequality (B.7) is true for all $t \geq 0$, it is also true that

$$\bar{F}_Y(e^{-1}(1+\epsilon)) \leq e^{-Nh^*(\theta)}, \tag{B.8}$$

where $h^*(\theta) = \sup_{t\geq 0} h(t,\theta)$.

It can now easily be shown that $h(t,\theta)$ is concave in t for fixed θ and that $h^*(\theta) = h(\theta/(1-\theta),\theta) = -\ln(1-\theta) - \theta > 0$ for all θ. Therefore, equation (B.5) becomes

$$M_N \leq 1 - [1 - (N+1)e^{-Nh^*(\theta)}](1 - e^{-1}(1+\epsilon)),$$

and, since $\lim_{N\to\infty}(N+1)e^{-Nh^*(\theta)} = 0$ for any θ,

$$\limsup_{N\to\infty} M_N \leq 1 - (1 - e^{-1}(1+\epsilon)),$$

for all $0 < \epsilon < e - 1$. By letting $\epsilon \to 0$ in the last inequality, it follows that

$$\limsup_{N\to\infty} M_N \leq 1 - (1 - e^{-1}) = e^{-1}. \tag{B.9}$$

Combining inequalities (B.4) and (B.9), it follows that

$$\lim_{N\to\infty} M_N = e^{-1}.$$

References

[1] E.D. Weinberger and S.A Kauffman, "The *NK* Model of Rugged Fitness Landscapes and its Application to Maturation of the Immune Response," *Journal of Theoretical Biology*, **141** (1989) 211–245.

[2] S. A. Kauffman, *The Origins of Order* (Oxford University Press, Oxford, 1993).

[3] B. Derrida, "Random-energy Model: An Exactly Solvable Model of Disordered Systems," *Physics Review B: Condensed Matter*, **24** (1981) 2613–2626.

[4] D. Solow, G. Vairaktarakis, S. Piderit and M. Tsai, "Managerial Insights into the Effects of Interactions on Replacing Members of a Team," (under revision for *Management Science*).

[5] D. A. Levinthal, "Adaptation on Rugged Landscapes," *Management Science*, **43** (1997) 934–950.

[6] E. D. Weinberger, "Local Properties of Kauffman's N-K Model: A Tunably Rugged Energy Landscape," *Phys. Rev. A: At. Mol. Opt. Phys.*, **44** (1991) 6399–6413.

[7] D. Solow, A. N. Burnetas, T. Roeder and N. S. Greenspan, "Evolutionary Consequences of Selected Locus-Specific Variations in Epistasis and Fitness Contribution in Kauffman's *NK* Model," *Journal of Theoretical Biology*, **196** (1999) 181–196.

[8] D. Solow, A. N. Burnetas, M. Tsai, and N. S. Greenspan, "Understanding and Attenuating the Complexity Catastrophe in Kauffman's *NK* Model of Genome Evolution," *Complexity*, **5(1)** (1999) 53–66.

[9] L. G. Mitten, "Preference Order Dynamic Programming," *Management Science*, **21** (1974) 43–46.

[10] C. Tovey, "Local Improvement on Discrete Structures," in *Local Search in Combinatorial Optimization*, edited by E. Aarts and J. K. Lenstra (J. Wiley, New York, 1997).

[11] S. Ross, *A First Course in Probability*, fifth edition, (Prentice Hall, Upper Saddle River, New Jersey, 1998).

[12] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications* (Jones and Bartlett, Boston, 1993).